TOKYO METROPOLITAN UNIVERSITY

Ross Recovery under the Tree Model and Its Application to Risk Management

by

Koji Anamizu

A thesis submitted in fulfillment for the
degree of Master of Finance

Department of Business Administration
Graduate School of Social Sciences

January 10th, 2018

Abstract

Ross (2015) has shown that real-world distributions can be derived from risk-neutral densities, named the Recovery Theorem, which makes the information embedded in option prices directly accessible to applications such as trading strategies, portfolio optimization and risk management. However, as we have to solve an ill-posed problem in the recovery process, application of the theorem to empirical problems is not straightforward. We propose a new method based on the trinomial tree model. Under the method, in addition to its accuracy and robustness, we can decrease the computational time drastically. We then apply the method to the stock and FX markets. By using the recovered real-world distribution, we create some early warning indicators to predict the future risk events, and show their effectiveness.

Keywords: Recovery theorem, physical distribution, tree model, risk management, early warning indicator, forward-looking approach

Acknowledgment

First, my deepest appreciation goes to Professor Masaaki Kijima of Master of Finance Program, Tokyo Metropolitan University. Without his guidance and persistent help, this thesis would not have been possible. His comments and suggestions were inestimable value for my study. I also have greatly benefited from Eri Nakamura and Hitomi Ito. As the member in the same laboratory, they always gave me constructive comments and also they checked my presentation paper. In addition, I appreciate my company's supervisors for giving me a chance to study at the university and their financial support. Finally, I would like to express my gratitude to my family for their moral support and warm encouragement.

Contents

Chapter 1

Introduction

The quality required in the risk management has changed significantly after the global financial crisis in 2008. Major Banks have been forced to develop new risk management tools to supplement the traditional risk measures such as Value at Risk (VaR). In this chapter, we explain the current environment of banks' risk management and popular risk-measures' drawbacks.

1.1 Risk Management after the Global Financial Crisis

The global financial crisis highlighted global regulatory weaknesses. After the crisis, a lot of regulations' changes were considered and actually executed by Basel Committee on Banking Supervision (BCBS) and other financial regulators such as Federal Reserve, European Banking Authority and Bank of England. The event showed us that there were too little liquidity and capital in the banking system. Furthermore, lacking foresight in risk management and delay of the loss recognition were recognized as one of the biggest issues among regulators and banks. The main reason why such problems happened was that banks had heavily depended only on the historical-based approach like VaR (i.e. backward-looking approach). Hence, they have proposed several methods in the field of risk management to incorporate the foreseeable future into their risks.

Implementation of the Forward-Looking Approach

For example, stress testing is becoming the major risk management tool to cover the traditional risk indicators' drawback. The main objective of stress testing is to capture the foreseeable scenario (i.e. forward-looking approach), which might be going to happen with high probability or affect negatively on their business. Before the crisis,

for supervisors, stress testing was ad hoc tool to assess the current state of their domestic banking industry in the aftermath of severe external shocks. Increasingly, however, they are becoming regular features of the ongoing regulatory process. Also, for banks, the importance of stress testing is increasing as well. Especially, banks operating internationally, so-called Global-Systemically-Important-Banks (G-SIBs), are facing more comprehensive stress testing requirements. Actually, current tests are placing greater emphasis on both qualitative and quantitative analysis to evaluate broader risks within banks, though initial stress testing exercises mainly focused only on quantitative indicators like the capital ratio.

In addition, in the accounting field, a new allowance calculation method, expected-credit-loss (ELS), was proposed in International Financial Reporting Standards (IFRS). Before 2008, incurred-loss-model served as the basis for accounting recognition and measurement of credit losses. Under incurred-loss-model, banks were required only certain period's expected loss calculated mainly by the historical loss data. On the other hand, ELS requires banks to take their future scenarios into account the calculation of the allowance. In general, ELS increases the amount of banks' allowance and also the implementation cost is expected to be very large because current banks' accounting systems need many improvements to deal with this new method. Therefore, this change is still a big issue among them.

Stress Testing

Using the United States' supervisory stress testing as the example, which is known as comprehensive capital analysis and review or CCAR, we explain the basic of stress testing framework.

CCAR is the United States regulatory framework introduced by the Federal Reserve to assess, regulate and supervise large banks. CCAR is executed once a year. The assessment is performed on both qualitative and quantitative basis. In the assessment on the quantitative basis, common-equity-tier-1 and tier-1-leverage-ratio are mainly used to check whether banks' capital structures are stable given the stress testing scenarios and also whether the planned capital distributions are viable and acceptable. On the qualitative basis, on the other hand, Federal Reserve recently focuses on wide range of topics such as methods of banks' internal risk management, model risks, and capital planning processes. If they fail to pass the minimum requirements, they are forced to change their next year's business plans, including dividend plans and investment strategies. Therefore, banks now spend a high amount of time and money on this program. Figure 1.1 is the process image of stress testing.

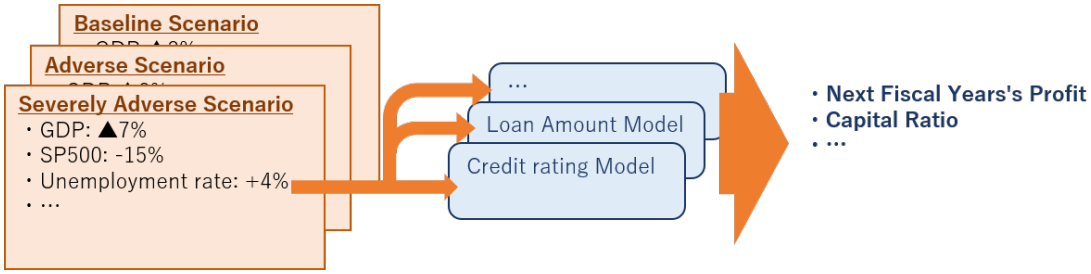


Figure 1.1: process image of stress testing

After preparing the data sets and developing each category's stress testing model, banks communicate internally and externally and decide the macro scenarios that impact significantly their business plans. Then they put the scenarios into the models to estimate the future profit, capital ratio and so on. In the case of CCAR, Federal Reserve investigates the result of the numerical test, its process, models, scenarios and current risk management framework to judge whether a bank passes the requirements or not.

1.2 Main Theme of the Thesis

Unfortunately, we can not say both stress testing and ELS are perfect forward-looking approaches. In most cases, the models used in the stress testing and ELS are based on the historical data. Therefore, when an event which has not occurred before happens, it is almost impossible to predict it beforehand. Moreover, scenarios used in the stress testing and ELS basically depend on economists' expert judgments or past historical events like Black Monday in 1987 and LTCM crisis in 1998, and also all foreseeable future scenarios are, of course, not included in stress testing.

To cover these drawbacks, banks are developing forward-looking indicators (i.e., early warning indicator (EWI)). Some of them are actually being used among practitioners. However, these indicators don't have robust theoretical backgrounds, and also their quality is generally not so high.

The aim of this thesis is to develop a better EWI not relying on the historical data. We consider to use the option data quoted on the market. The payoff of option is determined by the future distribution of underlying asset price and therefore the option prices contain the forward looking information. It can be expected that it is useful for predicting the future market condition. However, pricing the option value is done under risk-neutral measure which includes risk premium. So, we can't use the option data straightforwardly. Ross (2015) shows the real-world distributions can

be derived from risk-neutral densities, named the Recovery Theorem, which makes the information embedded in option prices directly accessible to applications such as trading strategies, portfolio optimization and risk management. In this thesis, by using the theorem, we recover the physical distribution and find the effective indicators that can predict the future market condition.

Structure of the Thesis

This thesis proceeds as follows. Chapter 2 summarizes the Recovery Theorem and how to apply it to real data. In addition, we also explain basic framework of the asset pricing theory as the reference. Moreover, we explain a problem the theorem has (ill-posed problem). Chapter 3 proposes a new approach (tree approach) to cope with the ill-posed problem. We also show its high accuracy and fast computation speed by comparing it with other approaches. Chapter 4 applies the tree approach to risk management. We apply it to S&P500 and USDJPY, and we recover their physical distributions. Then, we create EWI candidates and check each one's power for predicting the future risk event. In the final chapter, we conclude and describe future works.

Chapter 2

Asset Pricing Theory and the Recovery Theorem

In this chapter, we explain the basic framework of the asset pricing theory and the recovery theorem. After that, we introduce the literature related to the theorem.

2.1 Framework of the Asset Pricing Theory

In the framework of the asset pricing theory, a state price is an imaginary security price which brings a unit payoff when a certain state happens. This state price is determined uniquely under the assumption of an arbitrage-free and complete market. Define economic condition at time s as $\theta_s,\ s \in [t, T]$ and its state price at time T as $p(\theta_T|\theta_t)$. The time-t price of the security v_t can be expressed as

$$v_t = \int g(\theta_T) p(\theta_T|\theta_t) d\theta_T, \tag{2.1}$$

where $g(\theta_T)$ is a payoff function in the state θ_T. The sum of the state prices coincides with the zero-coupon bond price as

$$e^{-r(\theta_t)T} = \int p(\theta_T|\theta_t) d\theta_T, \tag{2.2}$$

where $r(\theta_t)$ describes a risk-free-rate function in the state θ_t. Normalized $p(\theta_T|\theta_t)$ satisfies the characteristics of the probability, and it is called the risk-neutral probability. By using (2.2), risk-neutral density $q(\theta_T|\theta_t)$ can be expressed as follows:

$$\begin{aligned} q(\theta_T|\theta_t) &\equiv \frac{p(\theta_T|\theta_t)}{\int p(\theta_T|\theta_t) d\theta_T}, & (2.3) \\ &= e^{r(\theta_t)T} p(\theta_T|\theta_t). & (2.4) \end{aligned}$$

Therefore, v_t can be rewritten as

$$v_t = \int e^{-r(\theta_T)T} g(\theta_T) q(\theta_T|\theta_t) d\theta_T, \tag{2.5}$$

$$= \mathbb{E}_t^{\mathbb{Q}} \left[e^{-r(\theta_T)T} g(\theta_T) \right]. \tag{2.6}$$

We similarly define the physical density $f(\theta_T|\theta_t)$. The pricing formula under physical measure can be expressed as below:

$$v_t = \mathbb{E}_t \left[e^{-r(\theta_T)T} g(\theta_T) \frac{q(\theta_T|\theta_t)}{f(\theta_T|\theta_t)} \right], \tag{2.7}$$

$$= \mathbb{E}_t[g(\theta_T) \underbrace{\frac{p(\theta_T|\theta_t)}{f(\theta_T|\theta_t)}}_{\equiv \phi(\theta_T|\theta_t)}], \tag{2.8}$$

$$\equiv \mathbb{E}_t[g(\theta_T)\phi(\theta_T|\theta_t)]. \tag{2.9}$$

where $\frac{q(\theta_T|\theta_t)}{f(\theta_T|\theta_t)}$ is a Radon–Nikodym derivative. $\phi(\theta_T|\theta_t)$ can be interpreted as a random variable which connects the security's payoff with the present value. It is called pricing kernel. In the Recovery Theorem, the pricing kernel and the physical distribution are simultaneously estimated uniquely under the condition that the state price is given.

2.2 Recovery Theorem

Assume that the economy has finite states in discrete time, and each state obeys the Markov process. The following formula is derived from (2.8).

$$p(\theta_{t+1}|\theta_t) = \phi(\theta_{t+1}|\theta_t) f(\theta_{t+1}|\theta_t). \tag{2.10}$$

Also, we give a specified function to describe $\phi(\theta_{t+1}|\theta_t)$ as

$$\phi(\theta_{t+1}|\theta_t) = \delta \frac{h(\theta_{t+1})}{h(\theta_t)}, \tag{2.11}$$

where δ is a fixed number, which can be assumed as a fixed discount factor, and $h(\theta)$ is a function of θ, which can be interpreted as an inverter's utility, according to Ross (2015).[1] Hence, (2.10) can be rewritten as

$$p(\theta_{t+1}|\theta_t) = \delta \frac{h(\theta_{t+1})}{h(\theta_t)} f(\theta_{t+1}|\theta_t), \tag{2.12}$$

[1] $h(\theta)$ is the same as the derivative of the consumption-based CAPM (CCPAM) utility function. Ross (2015), as the example that satisfies this assumption, introduces CCPAM.

or, in the matrix form,

$$\mathbf{DP} = \delta \mathbf{FD}. \tag{2.13}$$

In this case, $\mathbf{P}$ and $\mathbf{F}$ are $n \times n$ matrices, and $\mathbf{P} = (p_{ij})$ describes the state price transition matrix and $\mathbf{F} = (f_{ij})$ describes the physical probability transition matrix, respectively. $\mathbf{D}$ is an $n \times n$ diagonal matrix described as

$$\mathbf{D} \equiv diag(h(\theta = 1), \cdots, h(\theta = n)). \tag{2.14}$$

From now on, we assume that the function $h(\theta)$ does not depend on the time t, so we use θ instead of θ_t.

By solving (2.13) in $\mathbf{F}$, we get

$$\mathbf{F} = \left(\frac{1}{\delta}\right) \mathbf{DPD^{-1}}. \tag{2.15}$$

Since $\mathbf{F}$ is defined as a probability transition matrix, the sum of each row equals unity. Therefore, by defining $\mathbf{e} = (1, 1, \cdots, 1)^{\mathrm{T}}$, $\mathbf{F}$ satisfies

$$\mathbf{Fe} = \mathbf{e}. \tag{2.16}$$

Thus, equivalently, (2.16) can be expressed as

$$\begin{aligned} \mathbf{Fe} &= \left(\frac{1}{\delta}\right) \mathbf{DPD^{-1}e}, & (2.17) \\ &= \mathbf{e}. & (2.18) \end{aligned}$$

Hence, from (2.17) and (2.18), we get

$$\mathbf{PD^{-1}e} = \delta \mathbf{D^{-1}e}. \tag{2.19}$$

If we define $\mathbf{D^{-1}e}$ as

$$\mathbf{z} \equiv \mathbf{D^{-1}e}, \tag{2.20}$$

we have

$$\mathbf{Pz} = \delta \mathbf{z}. \tag{2.21}$$

(2.21) is interpreted as an eigenvalue problem. Since we assume that the market is arbitrage-free, $\mathbf{P}$ is a non-negative square matrix, and, by adding the assumption that

the matrix is a primitive matrix, we can apply the Perron–Frobenius theorem.[2] From the Perron-Frobenius theorem, the existence of a unique maximum eigenvalue which has a positive eigenvector is guaranteed. Therefore, we get $\mathbf{z}$ and δ. Thus, $\mathbf{D}$ can be calculated from $\mathbf{z}$. Finally, we get $\mathbf{F}$ from (2.15).

This process shows that we can estimate the pricing kernel and the physical probability transition matrix under the condition that only the state price transition matrix is given. This is the outline of the Recovery Theorem given in Ross (2015).

2.3 How to Apply to Empirical Data

The application procedure of the theorem to the empirical data is composed by 4 steps.

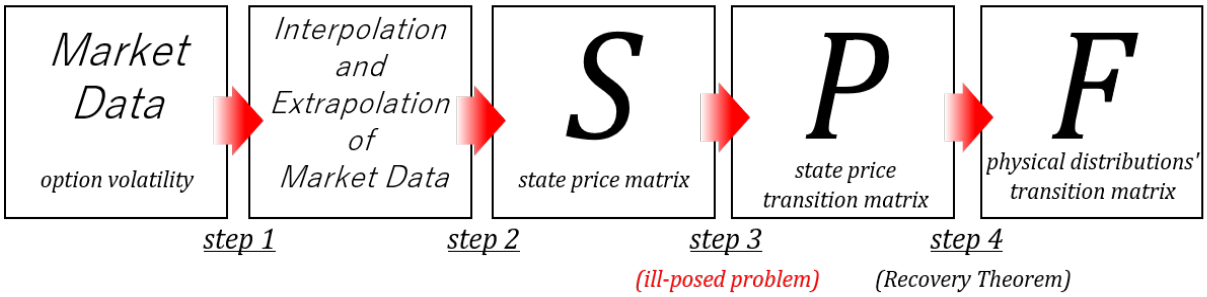


Figure 2.1: Application process of the Recovery Theorem

Step 1

This step is about the interpolation and the extrapolation of the implied-volatility data. In the process of application, option data (i.e., implied-volatility-surface) is required to create a state price matrix $\mathbf{S}$. However, not so many option implied-volatilities are observable in the market. (See Figure 2.2)

[2]Perron–Frobenius theorem is the theorem that guarantees following relations.
Let $\mathbf{A} = (a_{ij})$ be a primitive matrix. Then there exists an eigenvalue r such that:
(1) r is positive;
(2) the associated left and right eigenvectors can be chosen strictly positive componentwise;
(3) the eigen vectors associated with r are unique up to constant multiples;
(4) $r > |\lambda|$ for any other eigenvalue λ of $\mathbf{A}$;
(5) if $\mathbf{A} \geq \mathbf{B} \geq \mathbf{0}$ and β is an eigenvalue of $\mathbf{B}$, then$|\beta| \leq r$. Moreover, $|\beta| = r$ implies $\mathbf{B} = \mathbf{A}$;
(6) $\min_i \sum_j a_{ij} \leq r \leq \max_i \sum_j a_{ij}$.

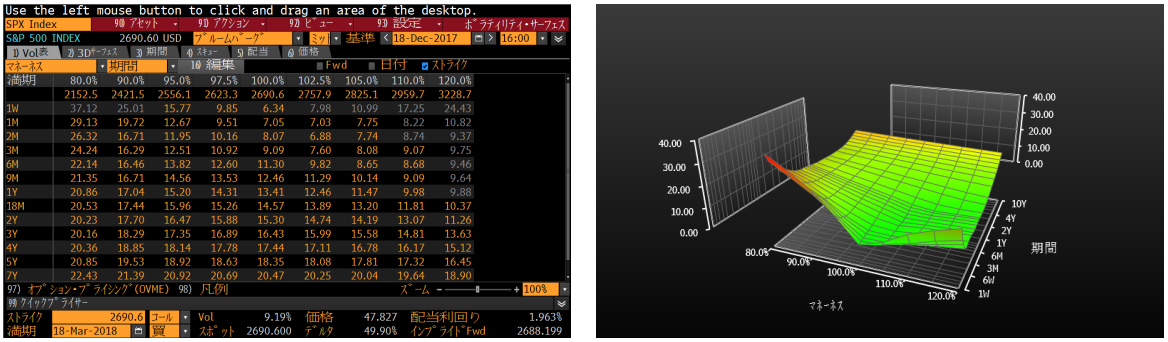


Figure 2.2: Volatility-surface for S&P500 European Option on December 18th, 2017 (Data source: Bloomberg L.P.)

To calculate the state price matrix in the next step, smooth implied-volatility-surface at each option maturity is required. We therefore interpolate the data. Also, around the deep out-of-the-money and deep in-the-money, option liquidity is so low that it is difficult to use it straightforwardly as the reliable data. An extrapolation is also necessary in this step.

In the field of interpolation and extrapolation, many approaches are proposed. We introduce some literature related to them. The first is Figlewsk (2010). Figlewsk (2010) proposes the interpolation and extrapolation approach, which uses spline functions and generalized-extreme-value (GEV) functions.

Spline Interpolation

The spline interpolation is a form of interpolation that is a special type of piecewise polynomial called a spline. Spline interpolation is often preferred over polynomial interpolation because the interpolation error can be made small even when using low degree polynomials for the spline. Generally, spline is the term for elastic rulers that are bent to pass through a number of predefined points ("knot"). The approach to mathematically model the shape of such elastic rulers fixed by $n+1$ knots $\{(x_i, y_i) : i = 0, 1, \cdots, n\}$ is to interpolate between all the pairs of knots (x_i, y_i) and (x_{i+1}, y_{i+1}) with polynomials $y = S_i(x), i = 1, 2, \cdots, n$.
As the spline will take a shape that minimizes the bending (under the constraint of passing through all knots), both y' and y'' will be continuous everywhere and at the knots. To achieve this, one must have that

$$S_i'(x_i) = S_{i-1}'(x_i) \quad i = 1, 2, \cdots, n, \tag{2.22}$$
$$S_i''(x_i) = S_{i-1}''(x_i) \quad i = 1, 2, \cdots, n. \tag{2.23}$$

This is the case of degree 3 spline function. Though the classical approach is to use polynomials of degree 3 (i.e. cubic-spline), Figlewsk (2010) proposes degree 4 spline function to get smooth risk-neutral distribution. Figlewsk (2010) also mentions the possibility of over-fitting in the case of degree 5 or higher spline functions.

Extrapolation of the Generalized-Extreme-Value Function

There are not so much implied volatility data quoted in the market, especially deep out-of-the-money and deep in-the-money. Thus, adding tail parts to the risk-neutral distribution is a necessary step, especially when we think about risk management. Figlewsk (2010) proposes one extrapolation approach based on GEV function. GEV function is defined as

$$F(x) = exp\left[-\left\{1+\xi\left(\frac{-x-\mu}{\sigma}\right)\right\}^{-1/\xi}\right], \tag{2.24}$$

where ξ is a fixed number, which is a parameter that controls the shape of the distribution, and μ and σ are parameters to set location and scale of the distribution. Hence, we need to set the three GEV parameters, which means that we prepare at least three constraint conditions on the tail. We use the expressions $F_{EVL}(\cdot)$ and $F_{EVR}(\cdot)$ to denote the approximating GEV distributions for the left and right tails, respectively, with $f_{EVL}(\cdot)$ and $f_{EVR}(\cdot)$ as the corresponding density functions. $F_{RND}(\cdot)$ and $f_{RND}(\cdot)$ denote the estimated empirical risk-neutral distribution and its density functions, respectivery.

Let $X(\alpha)$ denote the exercise price corresponding to the α-quantile of the risk-neutral distribution. That is, $F_{RND}(X(\alpha)) = \alpha$. First, we choose the value of α where the GEV tail is to begin, and then a second, more extreme point on the tail, that will be used in matching the GEV tail shape to that of the empirical risk-neutral density. These values will be denoted by α_{0R} and α_{1R}, respectively, for the right tail and α_{0L} and α_{1L} for the left.

After setting 4 points, α_{0L}, α_{1L} and α_{0R}, α_{1R}, consider to fit a GEV tail for the risk-neutral distribution. The first condition is that the total probability in the tail must be the same for the risk-neutral distribution and the GEV approximation. Figlewsk (2010) mentions that the GEV density has the same shape as the risk-neutral distribution in the area of the tail where the two distributions overlap, thus uses the other two degrees of freedom to set the two densities equal at α_{0R} and α_{1R}. Namely,

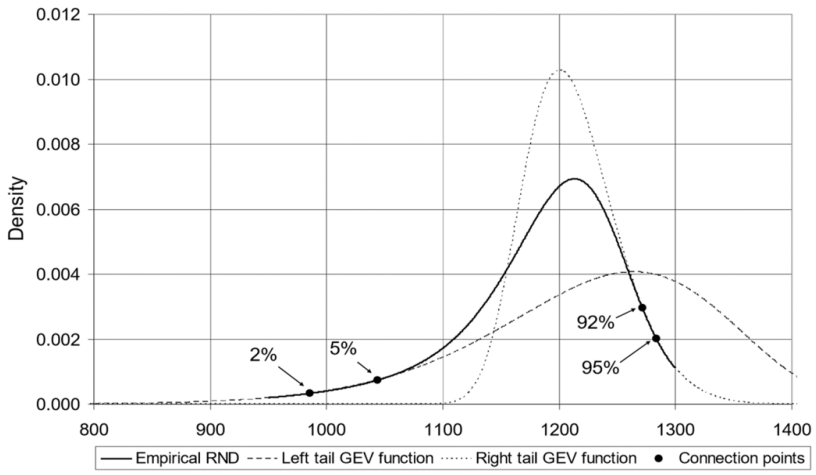


Figure 2.3: Image of the risk-neutral density and fitted GEV functions (Source: Figlewsk, S. (2010). "Estimating the implied risk neutral density for the US market portfolio." page 41)

$$F_{EVR}(X(\alpha_{0R})) = \alpha_{0R}, \tag{2.25}$$
$$f_{EVR}(X(\alpha_{0R})) = f_{RND}(X(\alpha_{0R})), \tag{2.26}$$
$$f_{EVR}(X(\alpha_{1R})) = f_{RND}(X(\alpha_{1R})). \tag{2.27}$$

Similarly, constraint conditions on the left tail can be described. Then, we solve each tail's minimization problem. The GEV parameters can be found easily by using a standard optimization procedure.

Interpolation and Extrapolation of the Double-Log-Normal Distribution

The second is Bliss and Panigirtzoglou (2002). They denote double-log-normal approximating function for interpolating and extrapolating the risk-neutral density.

Let $C_t(K)$ and $P_t(K)$ denote the European call and put option price at time t with strike K and option maturity T, respectively. $C_t(K)$ and $P_t(K)$ can be expressed by

$$C_t(K) = e^{-rT} \int_K^{+\infty} (S_T - K) dq(S_T), \tag{2.28}$$
$$P_t(K) = e^{-rT} \int_{-\infty}^{K} (K - S_T) dq(S_T). \tag{2.29}$$

Double-log-normal function $\hat{L}(\cdot)$ is denoted by using 2 log-normal distribution $L(\cdot)$ as

$$\hat{L}(S_T) = \theta L(S_T|\mu_1, \sigma_1, S_t) + (1-\theta)L(S_T|\mu_2, \sigma_2, S_t), \quad \theta \in [0,1], \tag{2.30}$$

$$L(S_T) = \frac{1}{S_T\sigma\sqrt{2\pi\tau}} exp\left\{\frac{-[logS_T - logS_t - (\mu - \frac{1}{2}\sigma^2)T]^2}{2\sigma^2 T}\right\}. \tag{2.31}$$

In the case of the European call option, by putting (2.30) into (2.28), we get

$$\begin{aligned}\hat{C}_t(K|\mu_1, \sigma_1, \mu_2, \sigma_2, \theta) = e^{-rT}\{\theta \int_K^{+\infty} (S_T - K)L(S_T|\mu_1, \sigma_1, S_t)dS_T \\ + (1-\theta)\int_K^{+\infty}(S_T - K)L(S_T|\mu_2, \sigma_2, S_t)dS_T\}.\end{aligned} \tag{2.32}$$

The European put option price can be described as well. 5 parameters, $\{\mu_1, \sigma_1, \mu_2, \sigma_2, \theta\}$, are estimated by solving the optimization problem (2.33) below, where N_c, N_p describe the number of call options and put options observed in the market. w_i, w_j are interpreted as weight for each optimization, and they are generally determined referring to each option's liquidity.

$$\begin{aligned}\min_{\mu_1,\sigma_1,\mu_2,\sigma_2,\theta} \quad & \sum_{i=1}^{N_c} w_i[C_t(K_i) - \hat{C}_t(K_i|\mu_1, \sigma_1, \mu_2, \sigma_2, \theta)]^2 + \\ & \sum_{j=1}^{N_p} w_j[P_t(K_i) - \hat{P}_t(K_i|\mu_1, \sigma_1, \mu_2, \sigma_2, \theta)]^2, \\ s.t \quad & \sum_{i=1}^{N_c} w_i + \sum_{j=1}^{N_p} w_j = 1, \quad w_i, w_j \geq 0.\end{aligned} \tag{2.33}$$

In Bliss and Panigirtzoglou (2002), they mention 2 log-normal functions are best to interpolate and extrapolate the distribution from the view point of the stability.

Step 2

Define $\mathbf{S}_t, (t = 1, \cdots, m)$ as a $1 \times n$ state price vector with option maturity t, and similarly define an $m \times n$ matrix $\mathbf{S}$ that contains each $\mathbf{S}_t, (t = 1, \cdots, m)$. The objective of step 2 is to estimate $\mathbf{S}$ from implied volatility data. The most famous method is based on Breeden and Litzenberger (1978). They show that we get the state price by differentiating the option price twice with the option's strike.[3] Many other methods, e.g., Bliss and Panigirtzoglou (2002), Melic and Thomas, Ludwig (2015) and Ludwig (2015), are also proposed.

[3]The derivation is contained in appendix A

Step 3

In Step 3, we estimate the state price transition matrix $\mathbf{P}$ from $\mathbf{S}$. There is no sophisticated method about estimating $\mathbf{P}$, so we need to think out the estimation method. Ross (2015) estimates $\mathbf{P}$ by using the optimization problem as

$$\max_{\mathbf{P}} ||\mathbf{S}_t\mathbf{P} - \mathbf{S}_{t+1}||_2^2, \tag{2.34}$$

where $||\cdot||_2^2$ denotes the Euclidean norm. Audrino, Huitema and Ludwig (2015) point out, in the case that the dimension of $\mathbf{P}$ is not small, the optimization problem in Ross (2015) causes the ill-posed problem. The ill-posed problem is a situation that there are some candidates of optimal solutions whose objective function values are almost the same. Thus, by adding the regularization term, Audrino, Huitema and Ludwig (2015) propose a new method, known as Tikhonov method, to curve the outbreak of the ill-posed problem:

$$\max_{\mathbf{P}} ||\mathbf{S}_t\mathbf{P} - \mathbf{S}_{t+1}|| + \varsigma||\mathbf{P}||_2^2, \tag{2.35}$$

where the second term is the regularization term, and ς is a regularization parameter, which controls the trade-off between the fitting and the stability of $\mathbf{P}$. In other words, in this method, each element of $\mathbf{P}$ can not reach a high number because of the regularization term.

In the same spirit, Kiriu and Hibiki (2015), by adding another type of the regularization term, propose a new approach:

$$\max_{\mathbf{P}} ||\mathbf{S}_t\mathbf{P} - \mathbf{S}_{t+1}|| + \varsigma||\mathbf{P} - \bar{\mathbf{P}}||_2^2, \tag{2.36}$$

where the second term is the regularization term like Audrino, Huitema and Ludwig (2015). The difference between them is the existence of $\bar{\mathbf{P}}$. Kiriu and Hibiki (2015), before solving the optimization problem, set $\bar{\mathbf{P}}$ that is expected to be similar to $\mathbf{P}$. This approach shows a more stable performance than Audrino, Huitema and Ludwig (2015).

However, this approach still has some drawbacks. First, we need to set ς and $\bar{\mathbf{P}}$ beforehand based on the historical data (i.e., backward-looking approach). In addition, how to set plausible $\bar{\mathbf{P}}$ is a quite difficult problem, and also the result fluctuates according to the level of ς.

There are other approaches proposed in the literature. For example, Fabio, Julian and Yang (2016) try to prevent the ill-posed problem by adding 6 constraint conditions including the single-peak-property. Morikawa (2016) solves the ill-posed problem by using the optimization problem with only the constraint condition of the single-peak property.

Step 4

Step 4 is straightforward. By using $\mathbf{P}$, simply apply the Recovery Theorem to recover a unique physical transition matrix $\mathbf{F}$.

2.4 Literature Review of the Recovery Theorem

The year when Steve Ross first released the theorem was 2011, though the year when the theorem was accepted by the Journal of Finance was 2015. After the first appearance, many researchers have been analyzing this theorem and trying to apply it to the empirical data.

There are 3 types of studies related to the Recovery Theorem. The first is the research how to solve the ill-posed problem as we have already mentioned.

The second is the numerical test by using the theorem. Ross (2015) applies it to S&P 500 to recover the physical distribution and compares it with the historical distribution. After that, Martin and Ross (2013) denote how to recover the interest rate distribution of the long-term government bond. Audrino, Huitema and Ludwig (2015) show S&P 500 physical distribution and calculate the difference between physical distributions and risk-neutral distributions (i.e., moment-risk-premium). Jensen, Lando and Pedersen (2015), by using S&P 500, research about the regression analysis for the distribution's expected return and the volatility under the physical measure.

The third is the expansion of the theorem. In the Recovery Theorem, discrete time and a finite number of states are assumed. Carr and Yu (2012) and Dubynskiy and Goldstein (2013) show the availability of application of the theorem under the continuous time assumption. In addition, Walden (2014) and Park (2014) propose the new theorem without the assumption of finite states. Also, Jensen, Lando and Pedersen (2015) point out that the assumption of the pricing kernel in Ross (2015) is not realistic intuitively and propose a new theorem without the assumption by introducing the term structure of the discount factor.

Chapter 3

The Tree Approach

Audrino, Huitema and Ludwig (2014), and Kiriu and Hibiki (2015) provide us stable state price transition matrix **P**. However, as we mentioned, it is quite difficult to explain the theoretical background of the regularization term. Moreover, these approaches use the historical data.

Fabio, Julian and Yang (2016) and Morikawa (2016) add some constraint condition to increase the stability. It decreases the degree of freedom of the optimization problem, their approaches therefore may archive stable result under the normal environment. However, it may not be able to fit the emergent situation such as the global financial crisis. This is the situation what risk managers really want to predict beforehand.

In this chapter, we propose a new approach, which is more stable and straightforward, and show the high accuracy of the recover and fast computational time.

3.1 Trinomial Tree Approach

3.1.1 Concept of the Approach

First, we describe the state price matrix for S&P 500 induced from the European option data to get the image that what type of a transition matrix is likely fitted to the data. In Figure 3.1 is the state price of the European option of S&P 500 on October 10th, 2010.

It describes that the state price is gradually diffusing as the option maturity goes on. Therefore, it seems possible to depict its transition matrix **P** by the tree structure. We name it "tree approach".

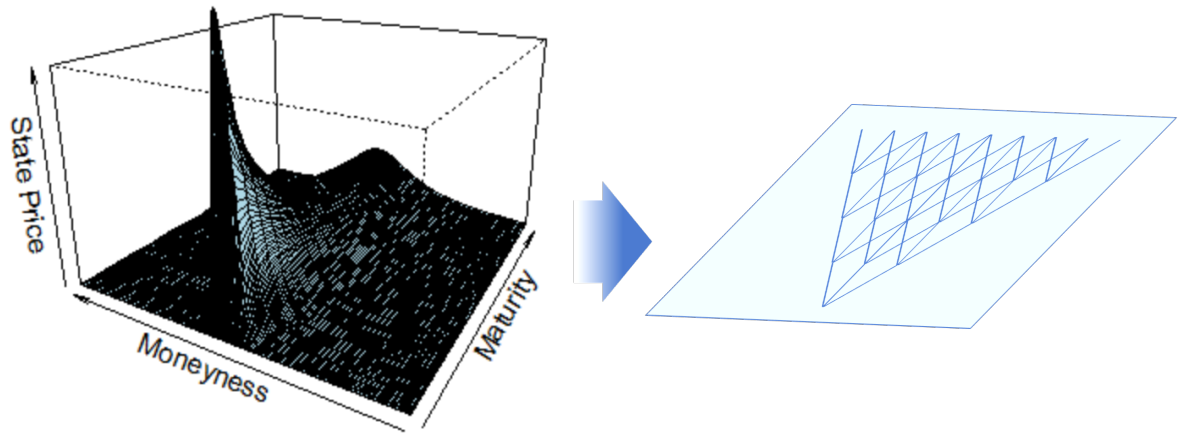


Figure 3.1: State price for S&P500 on October 10th, 2010 and trinomial tree structure

Theoretical background

We check the theoretical consistency of the tree approach. As the tree approach is an approximation method of describing a diffusion process, we confirm that the state price follows a stochastic differential equation.

Let $S(t)$ denote an asset price at time t and $B(t)$ denote a bank-saving-account at time t, and assume that they follow stochastic differential equations as follows:

$$\frac{dS(t)}{S(t)} = \mu(t)dt + \sigma(t)dz(t), \tag{3.1}$$

$$\frac{dB(t)}{B(t)} = r(t)dt. \tag{3.2}$$

Under the assumption of arbitrage-free and complete market,[1] let us define the exponential martingale as

$$Y(t) = e^{\theta z(t) - \theta^2 t/2}. \tag{3.3}$$

[1]The fundamental theorems of asset pricing provides necessary and sufficient conditions for a market to be arbitrage-free and for a market to be complete. In a discrete market, the following hold:

1. A discrete market, on a discrete probability space $(\Omega, \mathscr{F}, \mathscr{P})$, is arbitrage free if, and only if, there exists at least one risk-neutral probability measure that is equivalent to the original probability measure $\mathscr{P}$ (The First Fundamental Theorem of Asset Pricing).
2. An arbitrage free market (S, B) consisting of a collection of stocks S and a risk-free bond B is complete if and only if there exists a unique risk-neutral measure that is equivalent to $\mathscr{P}$ and has numeraire B (The Second Fundamental Theorem of Asset Pricing).

The above equation satisfies the following relation.

$$Q(A) = \mathbb{E}^{\mathbb{P}}[1_A Y(t)], \tag{3.4}$$

where Q denotes the risk-neutral probability. From (3.3), $Y(t)$ can be described as

$$dlogY(t) = \theta(t)dz(t) - \frac{1}{2}\theta^2(t)dt. \tag{3.5}$$

Therefore,

$$dlog\left(\frac{Y(t)}{B(t)}\right) = \left(-\frac{1}{2}\theta^2(t) - r(t)\right)dt + \theta(t)dz(t). \tag{3.6}$$

Since the state price density ϕ is described as $\frac{Y(t)}{B(t)}$, and under the arbitrage-free and complete market assumption, ϕ exists uniquely. Moreover, from Ito's lemma, ϕ is expressed as follows:

$$\frac{d\phi(t)}{\phi(t)} = -r(t)dt + \theta(t)dz(t). \tag{3.7}$$

Then, $\phi(t)$ and the state price $p(t)$ can be expressed as below:

$$\phi(t) = \frac{1}{e^{r(t)t}}\frac{d\mathbb{Q}(t)}{d\mathbb{P}(t)}, \tag{3.8}$$

$$p(t) = \phi(t)d\mathbb{P}(t). \tag{3.9}$$

We already showed that $\phi(t)$ can be described by a stochastic differential equation (i.e., diffusion process). So, if we assume $d\mathbb{P}$ follows a diffusion process, state price $p(t)$ also can be written as a diffusion process. Hence, it is natural to describe the state price transition matrix $\mathbf{P}$ by the tree structure.

Optimization Problem under the Approach

We can formulate the optimization problem under the tree approach as follows:

$$\min_{\mathbf{P}} \sum_{t=1}^{T-1} ||\mathbf{S}_t\mathbf{P}^k - \mathbf{S}_{t+1}||_2^2, \quad k \geq 1, \tag{3.10}$$

where $\mathbf{P}$ is a tri-diagonal matrix and k is a natural number, which controls the diffusion speed of the state price. It is much simpler than other approaches.

3.1.2 Numerical Test

We investigate the accuracy of the recovery under the tree approach by some numerical tests.

Procedure of Numerical Test

Figure 3.2 represents the framework of the test. We refer the method in Kiriu and Hibiki (2015).

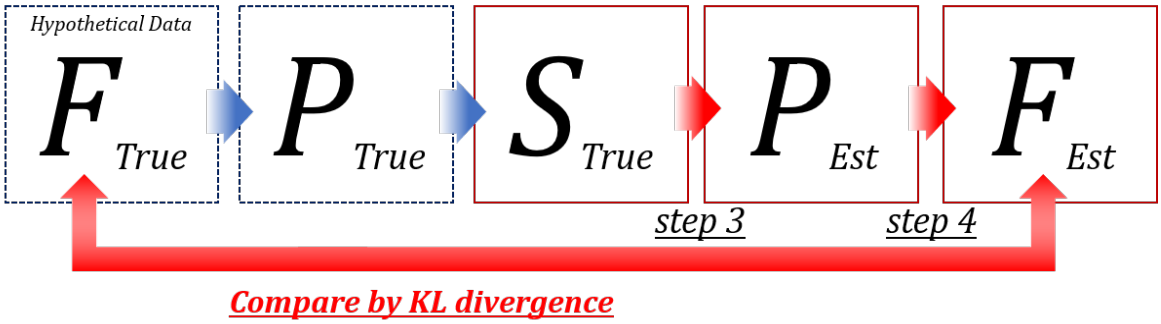


Figure 3.2: Procedure of the numerical test

1. Set the two simulated matrices; physical transition matrix $\mathbf{F}_{True}$ and pricing kernel $\mathbf{\Phi}_{True}$.
2. Calculate the true state price transition matrix $\mathbf{P}_{True}$ and true state price $\mathbf{S}_{True}$ in backward order.
3. Estimate $\mathbf{P}_{Est}$ from $\mathbf{S}_{True}$ by using Step 3 of the Recovery Theorem. $\mathbf{S}_{True}$ is composed by a large number of state price vector $\mathbf{S}_{t,True}$. In this step, we use 12 state price vectors $(\mathbf{S}_{Ture} = (\mathbf{S}_{1,True}, \cdots, \mathbf{S}_{12,True})^{\mathsf{T}})$
4. Similarly, Apply Step 4 of the theorem to recover $\mathbf{F}_{Est}$

Next, we describe the setting in more detail.

Physical Transition Matrix

The physical probability transition matrix $\mathbf{F}_{True}$ is a 13×13 matrix and it is discretely described by every 4%. So, this matrix is equally divided from -30% to 30%. Also, S&P 500 historical data (from January 3rd, 1950 to January 3rd, 2014) are used

to generate $\mathbf{F}_{True}$ because of the high liquidly and long data availability.

First, we set a reference date and calculate a 30-calendar-days return from the reference date. We use the day before a holiday in the case of a holiday. After counting the number of state transitions in each period, we sum up each element of the matrix. Finally, we divide each of them by each sum of the row elements to make a probability transition matrix.

Pricing Kernel

In the Recovery Theorem, we suppose that the pricing kernel is described by

$$\phi(\theta_{t+1}|\theta_t) = \delta \frac{h(\theta_{t+1})}{h(\theta_t)}. \tag{3.11}$$

As we explained, $h(\theta_t)$ is interpreted as the derivative, type of an investor's utility. In addition, we suppose that $h(\theta_t)$ is written as derivative of the constant relative risk aversion (CRRA) function $U(\theta_i) = \theta_i^{1-\gamma_R}/(1-\gamma_R)$, where γ_R is the relative risk aversion. If θ_i can be described by using the rate of return r_i, we can therefore calculate the pricing kernel as

$$\begin{aligned} p_{i,j}^{True} &= \delta \frac{U'(\theta_j)}{U'(\theta_i)} f_{i,j}^{True} && (3.12)\\ &= \delta \frac{\theta_j^{-\gamma_R}}{\theta_i^{-\gamma_R}} f_{i,j}^{True} && (3.13)\\ &= \delta \left(\frac{1+r_j}{1+r_i} \right)^{-\gamma_R} f_{i,j}^{True}. && (3.14) \end{aligned}$$

We therefore can calculate $\mathbf{P}_{True}$ from (3.14). In this test, $\gamma_R = 3$, and $\delta = 0.999$ are used.

Evaluation Criteria of the Estimation Accuracy

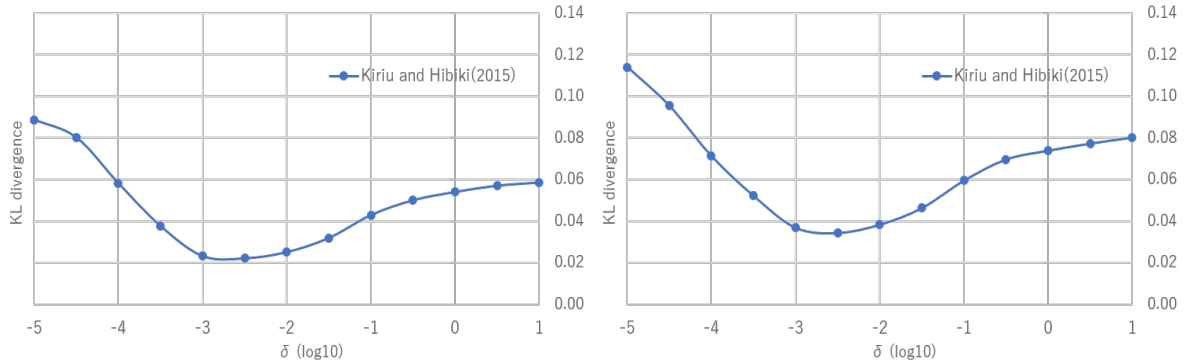
We use Kullback–Leibler divergence (KL divergence) as the evaluation criteria. Let us define F_i^{True} as the probability under the physical measure in the θ_i, and similarly define F_i^{Est} as well. After recovering $\mathbf{F}^{Est}$, compare $\mathbf{F}^{Est}$ and $\mathbf{F}^{True}$ as follows:

$$D_{KL}(\boldsymbol{F}^{Est}|\boldsymbol{F}^{True}) := \sum_{i=1}^{n} F_i^{Est} ln \left(\frac{F_i^{Est}}{F_i^{True}} \right). \tag{3.15}$$

When the estimated distribution is exactly equal to the true distribution, D_{KL} is equal to zero.

Result

First, we compare Kiriu and Hibiki (2015) approach with ς (ς=$10^{-5}, \cdots, 10^{1}$). Figure 3.3 shows the KL divergence of 3 month later and 6 month later. Around $\varsigma = 10^{-2.5}$, this approach attain the lowest KL divergence. However, the result fluctuates a lot when the level of ς moves.



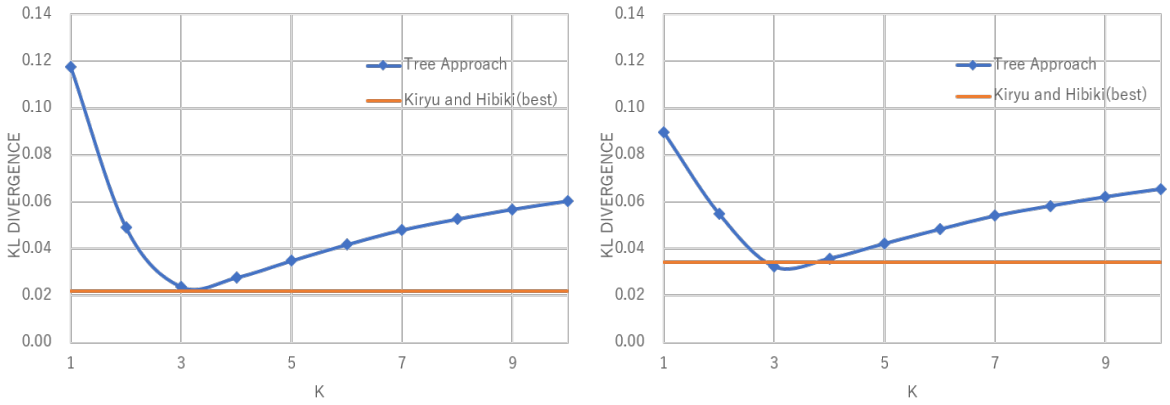
(a) 3 month later (b) 6 month later

Figure 3.3: KL divergence of the recovered physical distribution in the case of Kiriu and Hibiki (2015)

Next, Figure 3.4 shows the KL divergence in the case of the tree approach with k ($k = 1, \cdots, 10$). [2] The tree approach of both 3 month and 6 month later attains the lowest KL divergence in the case of $k = 3$. In addition, the result of the tree approach is as low as or a little bit better than best result of Kiriu and Hibiki (2015).

Same as Kirui and Hibiki (2015), parameter k influences the KL divergence. However, $\mathbf{P}$ is a tri-diagonal matrix in the tree approach. So, it is natural why the case of $k = 1$ marks the higher KL divergence. Therefore, $k = 1$ is generally never chosen. In addition, the cases of the higher k, $k = 6, \cdots, 10$, are almost meaningless, because $\mathbf{P}$ is a 13×13 tri-diagonal matrix. Therefore, in the case of $k = 5$, most elements in the matrix have positive number. For example, we show the $\mathbf{P}$ with $k = 2$ as below. Nearly 50% of the elements are covered even in the case of $k = 2$.

[2] In this analysis, we use a 13×13 matrix. So, higher k is not required. However, we are checking the higher case to see its behavior.



(a) 3 month later

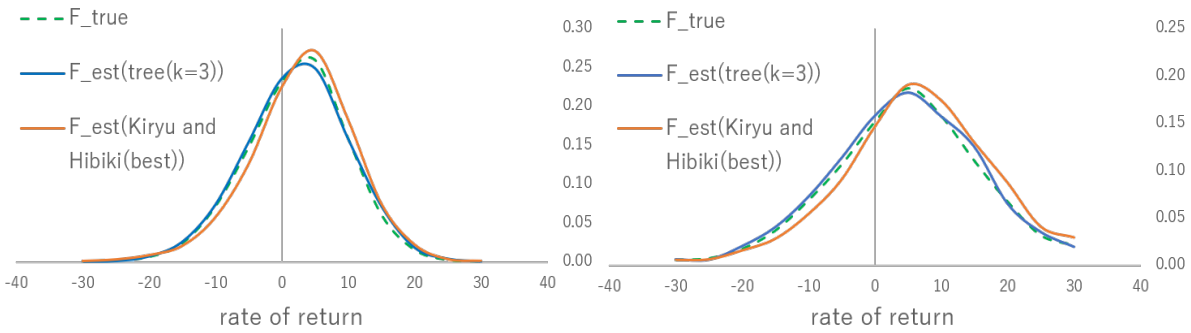
(b) 6 month later

Figure 3.4: KL divergence of the recovered distribution under the tree approach

$$\mathbf{P}^2 = \begin{pmatrix} 0.63 & 0.17 & 0.17 & 0.03 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.29 & 0.14 & 0.36 & 0.18 & 0.03 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.11 & 0.13 & 0.37 & 0.29 & 0.08 & 0.01 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.01 & 0.05 & 0.21 & 0.39 & 0.27 & 0.07 & 0.001 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.01 & 0.04 & 0.19 & 0.42 & 0.27 & 0.07 & 0.01 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.00 & 0.03 & 0.18 & 0.42 & 0.28 & 0.07 & 0.01 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.00 & 0.03 & 0.18 & 0.43 & 0.29 & 0.06 & 0.00 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.00 & 0.03 & 0.18 & 0.46 & 0.26 & 0.06 & 0.00 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.00 & 0.03 & 0.17 & 0.47 & 0.27 & 0.06 & 0.00 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.00 & 0.02 & 0.15 & 0.47 & 0.29 & 0.7 & 0.01 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.00 & 0.02 & 0.16 & 0.46 & 0.29 & 0.7 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.00 & 0.02 & 0.19 & 0.45 & 0.34 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.00 & 0.02 & 0.17 & 0.81 \end{pmatrix} \tag{3.16}$$

Moreover, it is understandable why the KL divergence increases as k increases. It seems very difficult to solve the optimization problem when k is higher, because the optimization problem becomes more complected.

Figure 3.5 shows recovered physical distributions of 3 month later and 6 month later. The new approach is more fitted to true physical distributions too.



(a) 3 month later

(b) 6 month later

Figure 3.5: Recovered distribution under tree approach

3.1.3 Characteristics of the Approach

We summarize the characteristics of the tree approach as bellow:

- < Robust theoretical background and not depending on historical data >
  In the case of Kiriu and Hibiki (2015), finding the suitable δ beforehand seems difficult and its theoretical background is very weak. However, the tree approach is correct that the state price transition can be depicted as a diffusion process, and it is consistent with finance theory. Furthermore, this approach is much stable with any k. Therefore, historical data is not needed at all.

- < Stable recovery >
  The recovery accuracy in the tree approach is better than Kiriu and Hibiki (2015).

- < Fast computational time of the recovery >
  The Recovery Theorem has been expected to apply to not only risk management, but also investment strategies. Therefore, computational time is a significant issue. As the elements that we have to estimate are much fewer in the tree approach, the recovery speed is faster. Actually, it takes only around 6.73 seconds to recover the 1 physical distribution in the tree approach. However, Kiriu and Hibiki (2015) need 19.60 seconds.

3.2 Tree Approach with Jumps

$\mathbf{P}^k$ induced from the tree approach is a multi-diagonal matrix. Except for upper and lower rows, each row has the same number of positive elements. So, if k is lower,

this approach can not catch the radical state price movement. In the tree approach, however, the KL divergence fluctuation is observable if higher k is used. In this section, in order to catch such radical state price movements, we consider adding the jump process to the tree approach. The processes described below are popular jump process, where J is a random variable, h is a fixed parameter, which satisfies $h = \mathbb{E}[J-1]$, A is an arbitrary financial instrument, which has drift μ_A and volatility σ_A. Also, N_t is a Poisson process with intensity λ. z_t, N_t and J are generally supposed to be respectively independent.

- Normal jump diffusion process

$$\frac{dA(t)}{A(t-)} = (\mu_A - \lambda h)dt + \sigma_A dz_t + (J-1)dN_t, \qquad J \sim \Phi(\mu, \sigma^2). \tag{3.17}$$

- Lognormal jump diffusion process (Merton model)

$$\frac{dA(t)}{A(t-)} = (\mu_A - \lambda h)dt + \sigma_A dz_t + (J-1)dN_t, \qquad lnJ \sim \Phi(\mu, \sigma^2). \tag{3.18}$$

- Laplacian jump diffusion process (Kou model)

$$\frac{dA(t)}{A(t-)} = (\mu_A - \lambda h)dt + \sigma_A dz_t + (J-1)dN_t. \tag{3.19}$$

  In the Kou model, the probability density function of lnJ is described as below. $\mathbb{1}$ is an indicator function.

$$p\eta_1 e^{-\eta_1 y}\mathbb{1}_{\{y\geq 0\}} + (1-p)\eta_1 e^{-\eta_2 y}\mathbb{1}_{\{y<0\}}, \qquad \eta_1 > 1, \eta_2 > 0. \tag{3.20}$$

Discrete compound Poisson distributions can be also implemented in the tree approach. However, we need to set each jump width beforehand. It means that it increases a chance of the using discretion. In addition, increasing the number of parameters eliminates the tree approach characteristics. Thus, in this thesis, we treat only continuous jump models.

3.2.1 Implementation Method

Same as the tree approach, the following optimization problem is used in the tree model with jumps:

$$\min_{\mathbf{P}} \sum_{t=1}^{n-1} ||\mathbf{S}_t\mathbf{P}^k - \mathbf{S}_{t+1}||_2^2. \tag{3.21}$$

In the tree approach, we assume that $\mathbf{P}$ is a tri-diagonal matrix. However, to take the jump process into account, we add a jump term in $\mathbf{P}$. For example, $\mathbf{P}$ in the tree approach with normal jump process is described as

$$
\begin{aligned}
P \;=\; & \underbrace{\begin{pmatrix}
p_{11} & p_{12} & & \cdots & \cdots & \\
p_{21} & p_{22} & p_{23} & & 0 & \\
 & \ddots & \ddots & \ddots & & \vdots \\
\vdots & 0 & & p_{n-1,n-2} & p_{n-1,n-1} & p_{n-1,n} \\
 & & \cdots & & p_{n,n-1} & p_{n,n}
\end{pmatrix}^{k}}_{Trinomial\ Tree\ Part} \\
+ \; & \underbrace{\begin{pmatrix}
\lambda\Delta t \int_{m_0} \phi_{\mu,\sigma}(s)ds & \lambda\Delta t \int_{m_1} \phi_{\mu,\sigma}(s)ds & & \cdots & \lambda\Delta t \int_{m_{n-1}} \phi_{\mu,\sigma}(s)ds \\
\lambda\Delta t \int_{m_{-1}} \phi_{\mu,\sigma}(s)ds & \lambda\Delta t \int_{m_0} \phi_{\mu,\sigma}(s)ds & & & \\
\vdots & & \ddots & \ddots & \vdots \\
\vdots & & & \lambda\Delta t \int_{m_0} \phi_{\mu,\sigma}(s)ds & \lambda\Delta t \int_{m_1} \phi_{\mu,\sigma}(s)ds \\
\lambda\Delta t \int_{m_{n-1}} \phi_{\mu,\sigma}(s)ds & & \cdots & \lambda\Delta t \int_{m_{-1}} \phi_{\mu,\sigma}(s)ds & \lambda\Delta t \int_{m_0} \phi_{\mu,\sigma}(s)ds
\end{pmatrix}}_{Jump\ Part},
\end{aligned}
\tag{3.22}
$$

where $\phi_{\mu,\sigma}(s)$ is a normal probability density function with mean μ and deviation σ, and m_i is the range.

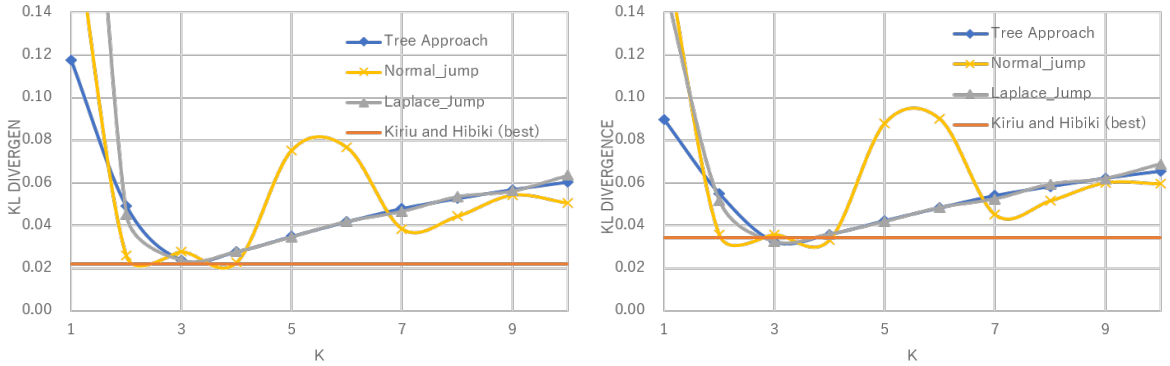
3.2.2 Numerical Test

We compare the KL divergence between the tree approach with jumps and the simple tree approach. Only normal model and the Kou model are used in this test. The method of this numerical test is the same as the method mentioned in 3.1.

Result

Figure 3.6 shows that the KL divergence with k of 3 months later and 6 months later, respectively. The KL divergences are as same as or worse than the simple tree approach. In addition, the result of the tree approach with normal jumps is fluctuated when k is higher. In the tree approach with jumps, the number of parameters is larger than the simple tree approach. Because of that, the optimization problem is more complicated and it might cause this fluctuation.

Under the tree approach with lower k, $\mathbf{P}^k$ can not describe the high volatility. So, at $k = 2$ or 3, the KL divergence of jump approaches are a little bit better than the tree



(a) 3 month later

(b) 6 month later

Figure 3.6: KL divergence in the tree approach with jump

approach. But, in the case of the tree approach with higher k, $\mathbf{P}^k$ can access to almost all states. So the difference between the tree approach and the tree approach with jumps is diminished. It might be useful to add a jump process to the tree approach when big matrix $\mathbf{P}$ is used. However, this result shows, in the case of smaller matrix such as 13×13, the merit of adding jump process is limited.

3.3 Non-Stationary Tree Approach

3.3.1 Drawback of the Tree Approach

As Figure 3.1 shows, the state price spreads to the left and right over the option maturity. Suppose that a state price with the option maturity t satisfies

$$dS(t) = \mu dt + \sigma dz \tag{3.23}$$

Then its volatility satisfies

$$\mathbb{E}[\{S(T) - \mathbb{E}[S(T)]\}^2] = \sigma^2 T. \tag{3.24}$$

Since each historical volatility can be calculated by using (3.24), we can create a daily single linear regression model for the volatility with the option maturity to get to know about the propensity of the volatility of the state price. Figure 3.7 describes the daily coefficients of the single regression models.

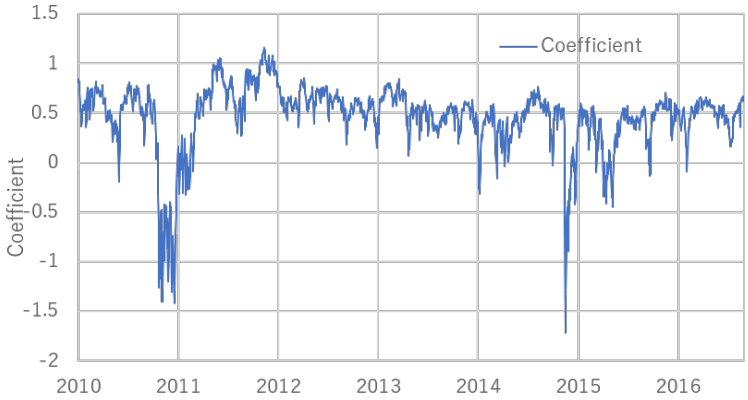


Figure 3.7: Coefficient of S&P500's single regression model (from October 20th, 2010 to June 1st, 2017)

Option maturity	Variance	Volatility
1 month	1.441	4.992
2 month	1.979	4.848
3 month	2.461	4.923
6 month	3.590	5.077
12 month	5.331	5.331
18 month	6.744	5.507
24 month	7.999	5.656

Table 3.1: Average state price volatility with the option maturity (from October 20th, 2010 to June 1st, 2017)

If coefficients are around 0, it means that the volatility is not fluctuated with the option maturity. Thus, we can assume that the volatility is a fixed number. However, coefficients of the historical data are actually fluctuated, and most of them attain positive numbers. In the tree approach, same $\mathbf{P}$ is used to all option maturities (i.e. stationary approach). Hence, we consider the non-stationary tree approach to grasp this propensity.

3.3.2 Implementation Method

To grasp the non-stationarity, we propose 2 methods.

1. Double tree approach
   This approach uses 2 state price matrices, $\mathbf{P}'$ and $\mathbf{P}''$, induced from 2 optimization

problems. $\mathbf{P}'$ denotes the transition matrix which priorities the first half of the data. $\mathbf{P}''$, on the other hand, is estimated with mainly the later half of the data. This procedure has 2 steps.

(a) Estimate $\mathbf{P}_t$ by solving the optimization problems $\mathbf{P}'$ and $\mathbf{P}''$:

$$\mathbf{P}_t = q_t \mathbf{P}'^k + (1 - q_t)\mathbf{P}''^k \quad 0 \leq q_t \leq 1, \tag{3.25}$$

$$\min_{\mathbf{P}'} \sum_{t=1}^{n-1} \omega'_t ||\mathbf{S}_t \mathbf{P}'^k - \mathbf{S}_{t+1}||_2^2, \tag{3.26}$$

$$\min_{\mathbf{P}''} \sum_{t=1}^{n-1} \omega''_t ||\mathbf{S}_t \mathbf{P}''^k - \mathbf{S}_{t+1}||_2^2, \tag{3.27}$$

where ω'_t, ω''_t are coefficient numbers calculated by functions such as power functions, $\omega'_t = a^{1-t}$, $\omega''_t = a^t$.

(b) Determine the optimal q_t in (3.25) by solving the optimization problem

$$\min_{q_t} \sum_{t=1}^{n-1} ||\mathbf{S}_t \mathbf{P}_t - \mathbf{S}_{t+1}||_2^2. \tag{3.28}$$

2. Tree approach with the option maturity t
This approach prepares $\mathbf{P}_t$ to each option maturity t. ω_t denotes a distribution like Laplacian distribution. After choosing the distribution, estimate $\mathbf{P}_t$ with each option maturity. Namely,

$$\min_{\mathbf{P}_t} \sum_{t=1}^{n-1} \omega_t ||\mathbf{S}_t \mathbf{P}_t - \mathbf{S}_{t+1}||. \tag{3.29}$$

3.3.3 Numerical Test

Procedure of Numerical Test

The basic procedure is the same as the numerical test in the tree approach. First, prepare the true physical transition matrix and the pricing kernel. Second, calculate the state price transition matrix and the state price. Finally, estimate the physical transition matrix by using the Recovery Theorem. All of the same parameters are basically used in this analysis. However, under the non-stationary assumption, in order to get more stable results, we use 48 state price vectors ($\mathbf{S}_{True} = (\mathbf{S}_{1,True}, \cdots, \mathbf{S}_{48,True})^\mathsf{T}$). Also, we use non-stationary $\mathbf{F}_{t,True}$ with each option maturity t instead of $\mathbf{F}_{True}$. Figure 3.8 shows the image of the procedure in the case of the double tree approach.

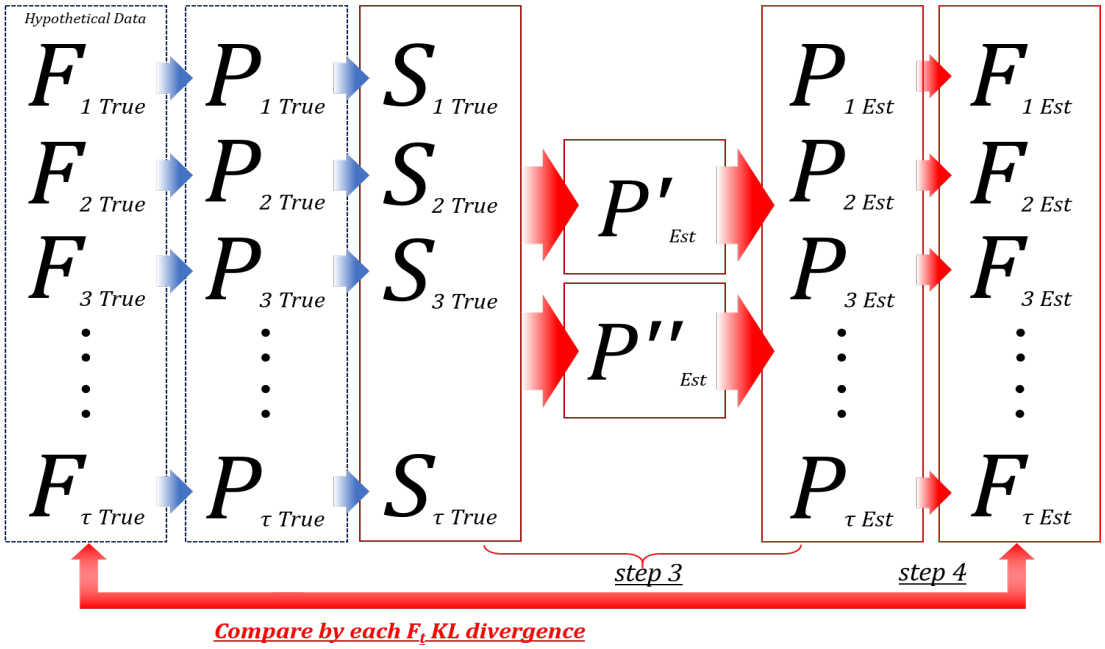


Figure 3.8: Procedure of the numerical test

Physical Transition Matrix

By adding non-stationary noise, we create $\mathbf{F}_{t,True}$ at each option maturity t. First, we denote a non-stationary noise. Suppose that a cumulative normal distribution is defined as Φ. $\phi_{\mu,\sigma}(\cdot)$ is a normal distribution with mean μ and voaltility σ.

$$\Phi_{\mu,\sigma}(x) \equiv \int_{-\infty}^{x} \phi_{\mu,\sigma}(s)ds \tag{3.30}$$

$$\begin{aligned} \therefore 1 = & \cdots + \int_{m_{-(n-1)}} \phi_{\mu,\sigma}(s)ds + \cdots + \int_{m_{-1}} \phi_{\mu,\sigma}(s)ds + \underbrace{\int_{m_0} \phi_{\mu,\sigma}(s)ds}_{(*)} \\ & + \int_{m_1} \phi_{\mu,\sigma}(s)ds + \cdots + \int_{m_{n-1}} \phi_{\mu,\sigma}(s)ds + \cdots \end{aligned} \tag{3.31}$$

Then, by using (3.31) as the noise, we create $\mathbf{F}_{t,True}$ as follows:

$$\mathbf{F}^{temp}_{t,True} \equiv \begin{pmatrix} f^{temp}_{t,11} & f^{temp}_{t,12} & & \cdots & \cdots & f^{temp}_{t,n,n} \\ f^{temp}_{t,21} & f^{temp}_{t,22} & f^{temp}_{t,23} & & & \\ & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & & f^{temp}_{t,n-1,n-2} & f^{temp}_{t,n-1,n-1} & f^{temp}_{t,n-1,n} \\ f^{temp}_{t,n,1} & & \cdots & & f^{temp}_{t,n,n-1} & f^{temp}_{t,n,n} \end{pmatrix} \tag{3.32}$$

$$\equiv \underbrace{\begin{pmatrix} f_{11} & f_{12} & & \cdots & \cdots & f_{n,n} \\ f_{21} & f_{22} & f_{23} & & & \\ & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & & f_{n-1,n-2} & f_{n-1,n-1} & f_{n-1,n} \\ f_{n,1} & & \cdots & & f_{n,n-1} & f_{n,n} \end{pmatrix}}_{F_{True}}$$

$$+ \quad 0.1 \times \frac{t}{12} \times \underbrace{\begin{pmatrix} 0 & \int_{m_1} \phi_{\mu,\sigma}(s)ds & & \cdots & \int_{m_{n-1}} \phi_{\mu,\sigma}(s)ds \\ \int_{m_{-1}} \phi_{\mu,\sigma}(s)ds & 0 & & & \\ \vdots & & \ddots & \ddots & \vdots \\ \vdots & & & 0 & \int_{m_1} \phi_{\mu,\sigma}(s)ds \\ \int^{m_{-(n-1)}} \phi_{\mu,\sigma}(s)ds & & \cdots & \int_{m_{-1}} \phi_{\mu,\sigma}(s)ds & 0 \end{pmatrix}}_{Non-Stationarity\ Noise} \tag{3.33}$$

To describe the propensity of the real-world volatility that the volatility increases according to the option maturity (See Table 3.1), $(*)$ in (3.31) is removed from (3.31) when adding it.

Finally, we normalize (3.33) to convert $\mathbf{F}^{temp}_{t,True}$ into a non-stationary probability transition matrix $\mathbf{F}^{n-s}_{t,True}$.

$$f^{n-s}_{t,i,j} \equiv \frac{f^{temp}_{t,i,j}}{\sum_{j=1}^{n} f^{temp}_{t,i,j}} \tag{3.34}$$

Approach for Estimating $\mathbf{F}_{\mathbf{est}}$

The double tree approach is used in this test. ω', ω'' are $0.3^t, 0.3^{1-t}$, respectively. Figure 3.9 shows that the shape of these power functions with the option maturity.

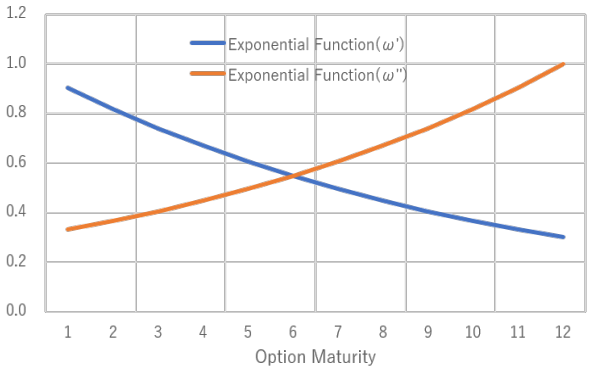


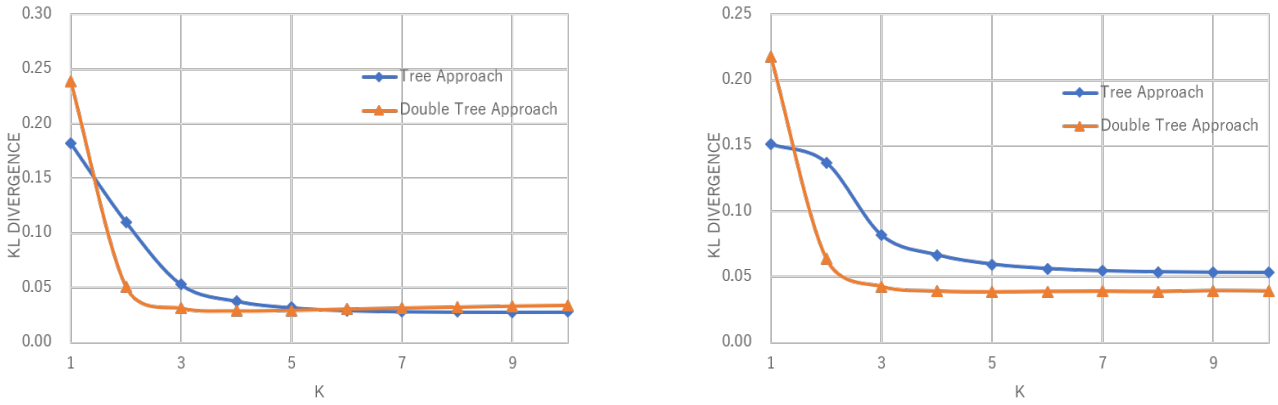
Figure 3.9: Power function ω_t', ω_t'' used in the analysis

Other Settings

The type of the pricing kernel that we assume is the same as 3.1.1. In addition, as the evaluation criteria, The KL divergence is used in the test as well.

Result

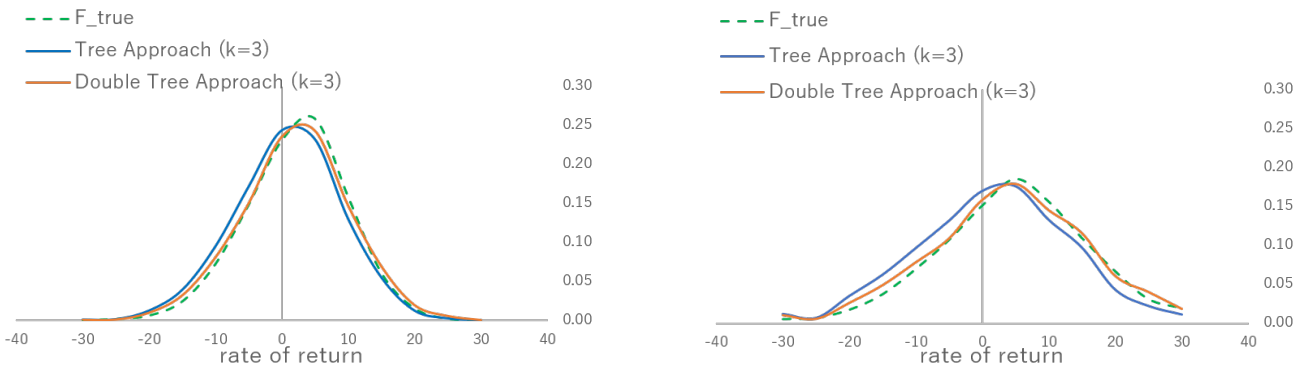
Figure 3.10 describes the KL divergence of both the double tree approach and the tree approach with k. For both 3 month and 6 month results, the non-stationary tree approach marks the better KL divergence than the tree approach in most of the case. Since the optimization problems are more complicated as k increases, especially when we use the double tree approach, the KL divergence of the stationary tree approach is better than the result of the double tree approach when k is higher than 6 in the 3 month result. However, as we explained, results with higher k is not significant.



(a) 3 month later

(b) 6 month later

Figure 3.10: KL divergence of the recovered distribution under the double tree approach



(a) 3 month later

(b) 6 month later

Figure 3.11: Recovered distribution under the double tree approach

Figure 3.11 compares the recovered distributions estimated from the double tree approach with the tree approach. Recovered distributions by the double tree approach are fitted to $F^{n-s}_{t,True}$ more accurately than distributions by the tree approach.

However, the double tree approach has some drawbacks. A lot of state price vectors are required to solve the stable optimization problems when we estimate $\mathbf{P}'$ and $\mathbf{P}''$. As we use power functions, we can not use all state price data. Moreover, the computational time is larger than the tree approach because of the increase of the number of optimization problems. Hence, it is better that we use both tree approaches for different purposes in accordance to the situation.

Chapter 4

Application to Risk Management

In this chapter, we apply the Recovery Theorem with the tree approach to risk management. The main aim of this analysis is to find effective EWIs (early warning indicators) for predicting the serious financial shocks such as the global financial crisis. First, we recover the physical distribution from S&P 500 historical data. Then, from the result, we create some indicators which seem to have a power for predicting the future distribution. Finally, we check their effectiveness by backtesting. The result shows that some indicators seem to be able to predict the future risk events. In addition, in the case of the foreign exchange rate USDJPY, we get almost the same result.

4.1 Analysis of S&P 500

4.1.1 Procedure

Figure 4.1 represents the framework of the EWI analysis. The first half of the procedure is the ordinal steps of the Recovery Theorem. After getting the option volatility data, interpolate and extrapolate it and recover the physical distribution.

The latter half is the new step for analyzing the effectiveness of the EWI. Through these steps, we create some EWI candidates and check their power for predicting the future serious risk events. In step D, set some indicators which seem to be effective for predicting the 1 month later event. These indicators are created by the recovered physical distribution $\mathbf{F}_{dist}$ and the risk-neutral distribution $\mathbf{Q}_{dist}$. In step E, define the risk event for backtesting. In step F, we do backtesting and investigate how accurately EWIs predict the future risk events.

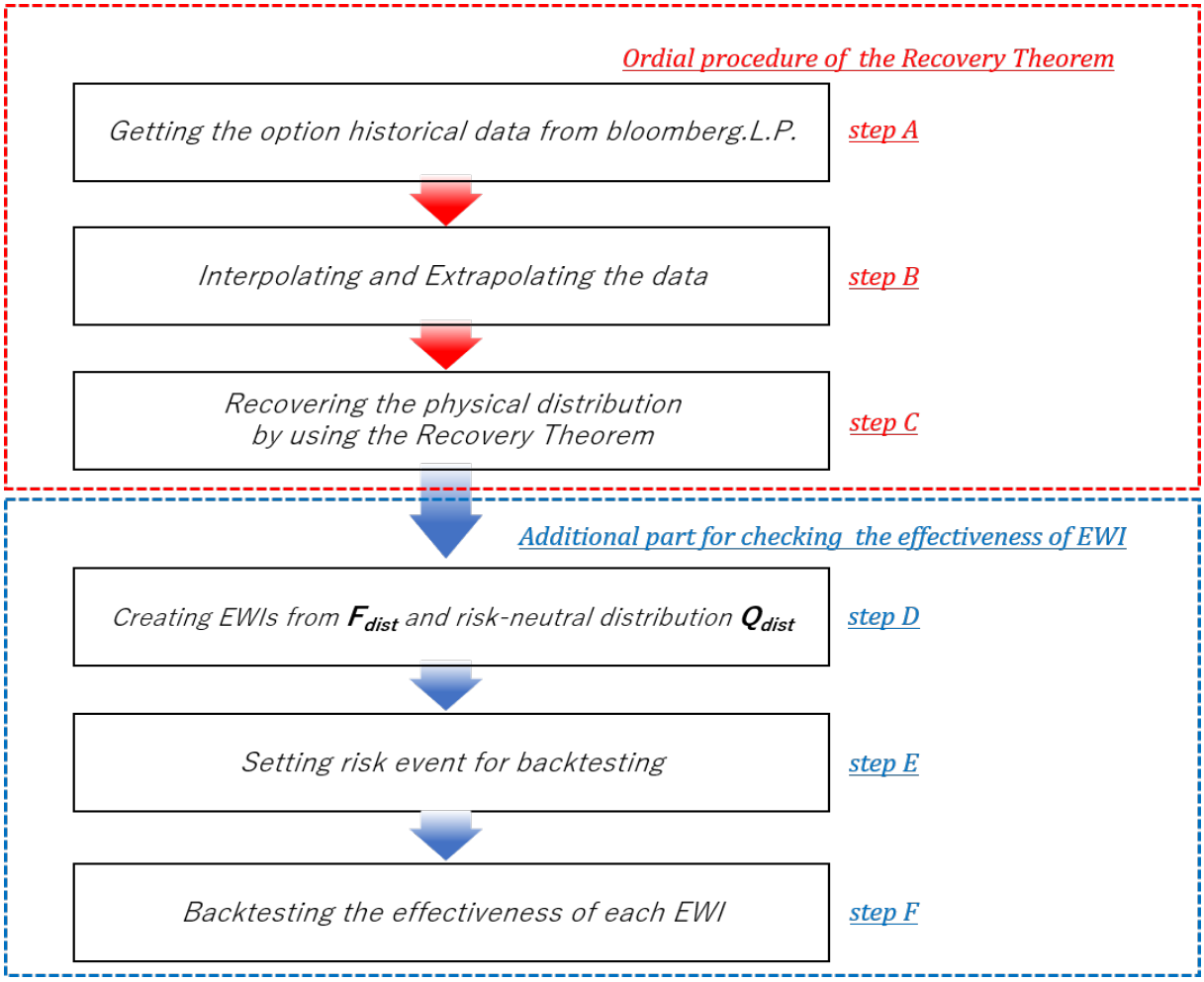


Figure 4.1: The image of the procedure

4.1.2 Setting

Data

In this analysis, data from Bloomberg L.P is used. We chose the option data which has the most detailed moneyness classifications and sufficient option maturities. The data covers the wide range of moneyness, {80, 90, 95, 97.5, 100, 102.5, 105, 110, 120 }, and it is available from June 24, 2005. In the analysis, we use 3005 daily data from June 24th, 2005, to June 1st, 2017. It has a wide range of option maturities too, {30day, 60day, 90day, 180day, 360day, 540day, 720day }. However, we avoid using {540d, 720d } because of their low liquidity in the market.

Approach for estimating P

The stationary trinomial tree approach that we developed is used in this analysis. Parameter $k = 4$ is used because of well understandability. Since we are predicting 1 month later distribution, $\mathbf{P}^k$ is a 1 month transition matrix. In the case $k = 4$, $\mathbf{P}$ itself is interpreted as an 1 week transition matrix. In addition, in this analysis, $\mathbf{P}$ is a 13×13 matrix.

Method of interpolation and extrapolation

Degree 4 spline function is used to interpolate the volatility curve in the direction of the moneyness. In addition, we similarly use 3 spline function for interpolating in the direction of the option maturity. Furthermore, we use GEV function for adding the tail part of the risk-neutral distribution calculated from the implied-volatility data.

EWI candidates

The moment, which are calculated from the recovered physical distribution, and the moment-risk-premium, which is the difference between the moment of the risk-neutral distribution and the moment of the recovered physical distribution, are mainly used as the source of WEI candidates. Each moment (mean: μ, variance: σ^2, skewness: ν, kurtosis: ξ) is calculated as follows:

$$\mu = E[X] = \int_{-\infty}^{\infty} X dF(X) \approx \sum_{i=1}^{N} X_i f(X_i) \Delta X_i, \tag{4.1}$$

$$\sigma^2 = E[(X-\mu)^2] \approx \sum_{i=1}^{N} (X_i - \mu)^2 f(X_i) \Delta X_i, \tag{4.2}$$

$$\nu = \frac{E[(X-\mu)^3]}{\sigma^3} \approx \frac{1}{\sigma^3} \sum_{i=1}^{N} (X_i - \mu)^3 f(X_i) \Delta X_i, \tag{4.3}$$

$$\xi = \frac{E[(X-\mu)^4]}{\sigma^4} \approx \frac{1}{\sigma^4} \sum_{i=1}^{N} (X_i - \mu)^4 f(X_i) \Delta X_i, \tag{4.4}$$

where X_i is the rate of return at state i, $f(X_i)$ is a probability distribution function, and since the economic states are divided into 13, N is 13.

In addition, we also calculate Semi-Variance (SV), ES and VaR as other source of

EWI candidates. SV and ES are calculated as

$$SV = E[|X-\mu|_-^2] \approx \sum_{i=1}^{N} |X_i-\mu|_-^2 f(X_i)\Delta X_i, \tag{4.5}$$

$$ES(\alpha) = \int_{-\infty}^{\alpha} X dF(X) \approx \sum_{i=1}^{n} X_i f(X_i)\Delta X_i, \tag{4.6}$$

$$\tag{4.7}$$

where α is a percentile point of ES and n is a nonnegative number which satisfies $\sum_{i=1}^{n} f(X_i)\Delta X_i = \alpha$. In addtion, VaR($\alpha$) is determined by the linear combination of 2 points which are next to α.

After calculating moments, moments-risk-premiums, SV, ES and VaR, we create 261 EWI candidates in total based on the concepts as below:

- $< Pattern1 > \cdots$ Capturing the radical change of the distribution
  The average of 1, 2 or 3 consecutive days' moment, moment-risk-premium, SV, ES or VaR is out of the range (75% or 90%) of its own distribution based on the latest business 20days.

- $< Pattern2 > \cdots$ Capturing the trend of the distribution change
  The average of 1, 2 or 3 consecutive days' moment, moment risk premium, SV, ES or VaR changes continuously(or with high frequency) in 3(4 or 5) consecutive days.

Risk Event

We first investigate the market risk calculation method defined by BCBS to set the reasonable risk event.

In January 2016, a new capital-requirements-rule in market risk was released from BCBS. Though it will not be effective to international banks until 2019, it is expected that it will be the new standard for market risk management.

In the document, international banks are required to measure the capital requirement related to market risk based on the risk factors' liquidity horizon.

Risk factor category	n
Interest rate: specified currencies - EUR, USD, GBP, AUD, JPY, SEK, CAD and domestic currency of a bank	10
Interest rate: – unspecified currencies	20
Interest rate: volatility	60
Interest rate: other types	60
Credit spread: sovereign (IG)	20
Credit spread: sovereign (HY)	40
Credit spread: corporate (IG)	40
Credit spread: corporate (HY)	60
Credit spread: volatility	120
Credit spread: other types	120
Equity price (large cap)	10
Equity price (small cap)	20
Equity price (large cap): volatility	20
Equity price (small cap): volatility	60
Equity: other types	60
FX rate: specified currency pairs[37]	10
FX rate: currency pairs	20
FX: volatility	40
FX: other types	40
Energy and carbon emissions trading price	20
Precious metals and non-ferrous metals price	20
Other commodities price	60
Energy and carbon emissions trading price: volatility	60
Precious metals and non-ferrous metals price: volatility	60
Other commodities price: volatility	120
Commodity: other types	120

Figure 4.2: Liquidity horizon of risk factors (Source: Bank for international settlements (2016). "Minimum capital requirements for market risk." page 55.)

Figure 4.2[37] shows each risk factor's liquidity horizon. Banks will have flexibility in devising the precise nature of their models, but the minimum capital requirement will apply for the purpose of a floor of the capital charge. In that process, ES for a liquidity horizon must be calculated from ES at a base liquidity horizon of 10 days with scaling.

In general, it is considered that rules made by BCBS are conservative. S&P 500 is based on the market capitalizations of 500 large companies having common stocks listed on the NYSE or NASDAQ. Intuitively Figure 4.2 is too conservative too, because it takes 10 business days to sell such major stocks in the market. Therefore, S&P 500 drop in cumulative 5 business days is used in this analysis. Finally, we set "over 10% drop in cumulative 5 business days within 20 business days" as the risk event.

[37]The footnote [37] of Figure 4.2 is as follows.
USD/EUR, USD/JPY, USD/GBP, USD/AUD, USD/CAD, USD/CHF, USD/MXN, USD/CNY, USD/NZD, USD/RUB, USD/HKD, USD/SGD, USD/TRY, USD/KRW, USD/SEK, USD/ZAR, USD/INR, USD/NOK, USD/BRL, EUR/JPY, EUR/GBP, EUR/CHF and JPY/AUD.

Measure for assessing EWI effectiveness

We assess EWIs performance based on 2 measures as below. Measure 1 is mainly used trough the analysis.

$$< Measure\ 1 > \quad : \quad \frac{\sum_t \mathbb{1}_{\{Risk\ Event\ happens\ from\ t\ to\ t+20\ while\ EWI\ alarms\ at\ t\}}}{\sum_t \mathbb{1}_{\{EWI\ alarms\ at\ t\}}}$$

$$< Measure\ 2 > \quad : \quad \frac{\sum_t \mathbb{1}_{\{EWI\ alarms\ at\ t\ while\ Risk\ Event\ happens\ from\ t\ to\ t+20\}}}{\sum_t \mathbb{1}_{\{Risk\ Event\ happens\ from\ t\ to\ t+20\}}}$$

4.1.3 Backtesting

First, we show the transition of moments, VaR(75%), ES(75%), and SV under both the physical measure and risk-neutral measure.

Moment, VaR, ES and Semi-Variance

Following characteristics are recognized from the figures.

- The mean under the risk-neutral measure moves around 1% stably. This result is almost consistent that all drifts under risk-neutral measure are as same as the risk-free rate under the risk-neutral valuation. On the other hand, the mean under the physical measure fluctuates with downward direction, especially around the global financial crisis and the European sovereign debt crisis. It enables us to get a aspect that market participants expected negative rate of return during such crisis. In addition, most of the time, positive risk-premium is observable.

- In the case of the variance, a high correlation is observable between the risk-neutral and the physical measure. Also, their levels are almost the same in entire period, though the variance of the physical distribution fluctuates more. It is also consistent from the viewpoint of the measure change.

- Like results of the mean and the variance, the risk-neutral distribution is more stable both in the skewness and the kurtosis. The kurtosis, however, seems to be more stable than the skewness. Furthermore, a positive correlation exists in the case of the skewness.

- The result of the SV is very similar to the variance. The risk-neutral and the physical measure have a high positive correlation.

- In the case of the VaR and the ES, their lines are very stable under the risk-neutral measure. In the case of the physical measure, on the other hand, the VaR and the ES move adversely and attain the lowest level in the financial global crisis and

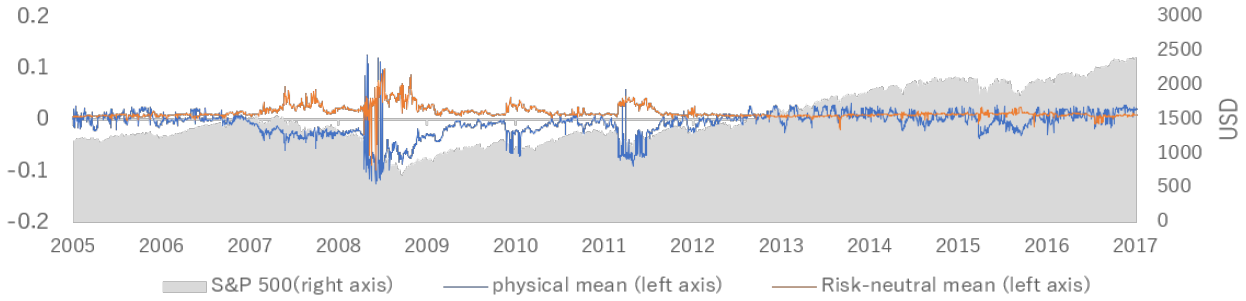


Figure 4.3: mean of the recovered physical distribution

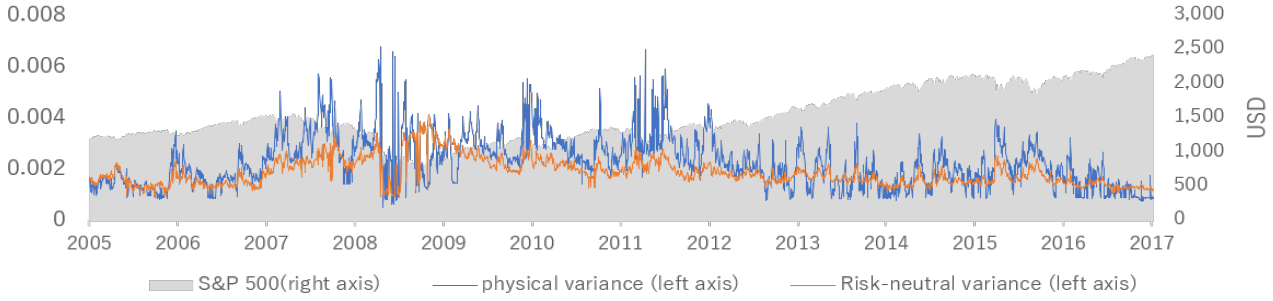


Figure 4.4: deviation of the recovered physical distribution

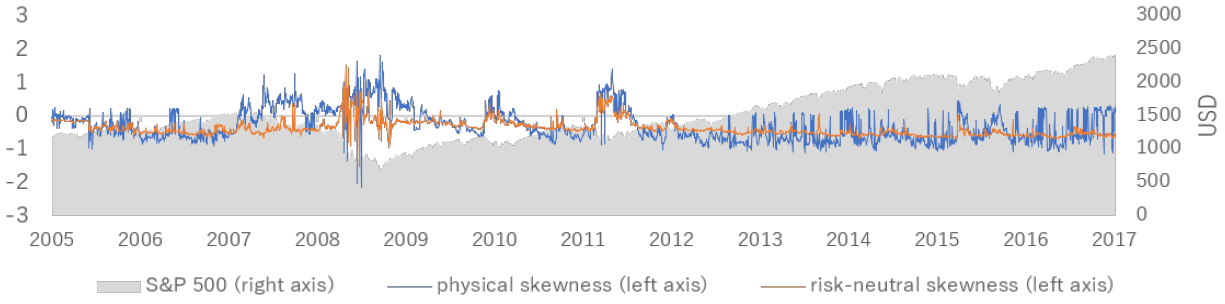


Figure 4.5: skewness of the recovered physical distribution

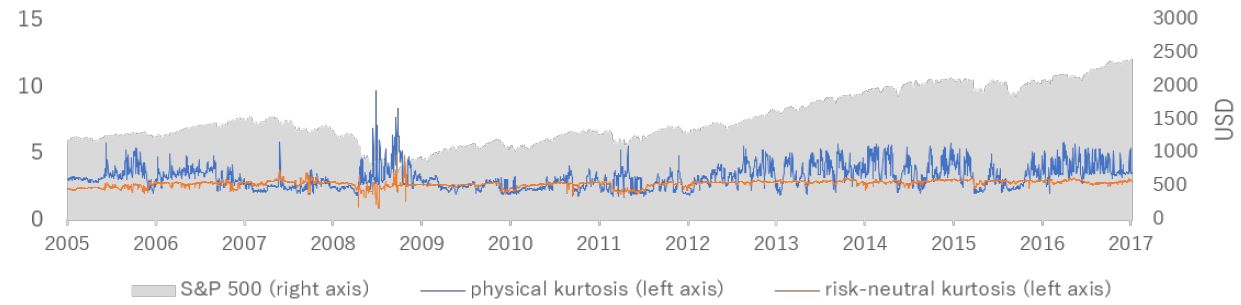


Figure 4.6: kurtosis of the recovered physical distribution

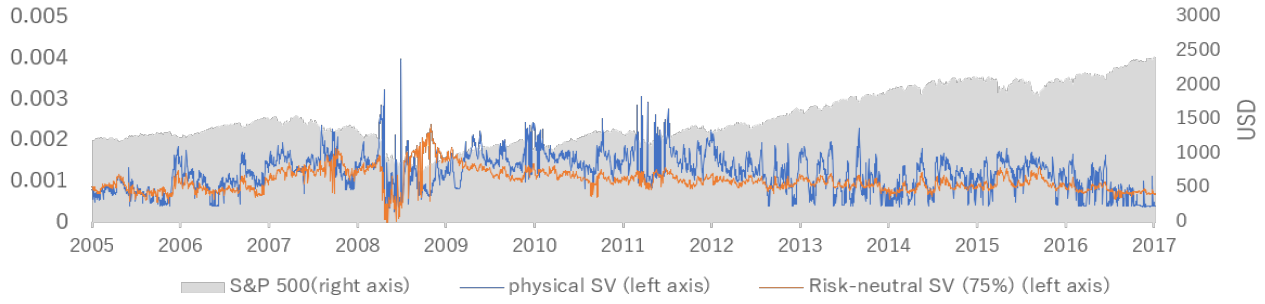


Figure 4.7: SV of the recovered physical distribution

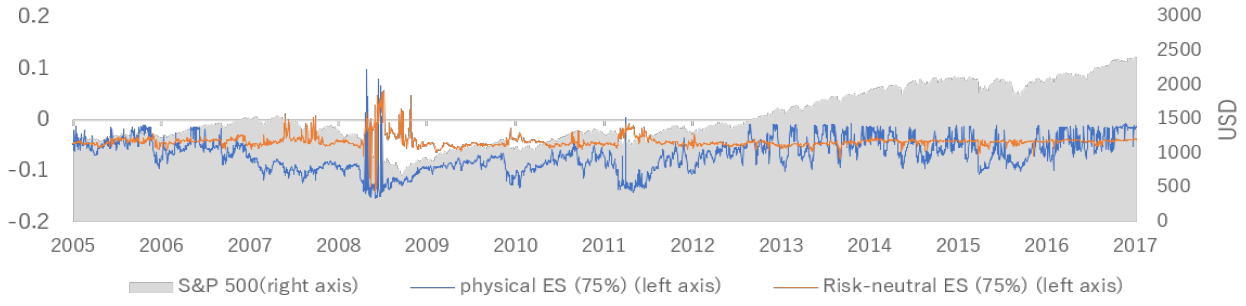


Figure 4.8: ES (75%) of the recovered physical distribution

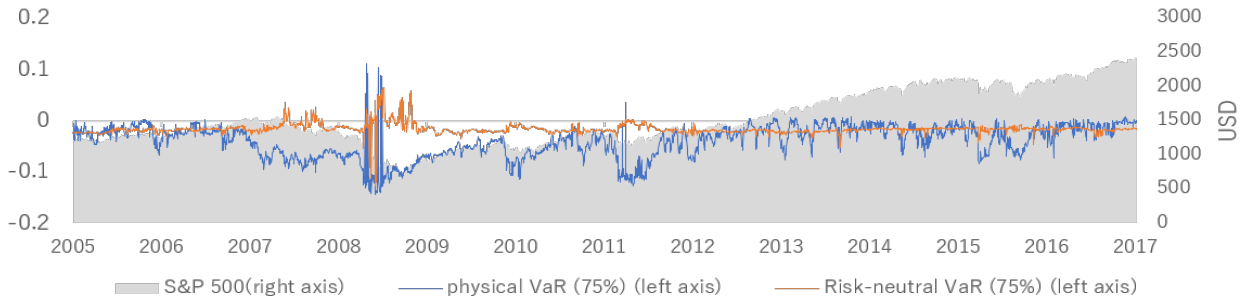


Figure 4.9: VaR (75%) of the recovered physical distribution

the European sovereign debt crisis. So It can be known that market participants had negative expectations during such big crisis.

Next, we show the correlation between each factor. Table 4.1 is a Pearson correlation matrix of moments, SV, ES(75%) and VaR(75%). Also, Table 4.2 shows the case of risk-premium-moments.

	Mean	Variance	Skewness	Kurtosis	SV	ES(75%)	VaR(75%)
Mean	1.000	0.4823	0.0891	0.0271	0.4252	0.6352	0.7625
Variance		1.000	0.1885	0.0916	0.9788	-0.3097	-0.1861
Skewness			1.000	-0.0208	0.1266	0.0126	-0.0552
Kurtosis				1.000	0.0985	-0.1039	-0.0552
SV					1.000	-0.4082	-0.2499
ES(75%)						1.000	0.9582
VaR(75%)							1.000

Table 4.1: Pearson correlation matrix of the physical measure

Correlation Between the Actual Return and the Recovered Return

The mean of the physical distribution is the most important factor among all factors. In order to test the effectiveness of this factor, we compare actual S&P 500 and estimated S&P 500 calculated by the mean under the recovered physical distribution 20days before. In general, it is known that the option market participants' data is ahead of the actual data. However, estimated S&P 500 induced from mean seems to be delayed in contrast to actual S&P 500 (See Figure 4.10). Therefore, we check the

	Mean	Variance	Skewness	Kurtosis	SV	ES(75%)	VaR(75%)
Mean	1.000	0.1713	-0.1355	0.0194	0.3027	0.8850	0.9467
Variance		1.000	-0.0630	0.0941	0.9473	-0.2313	-0.1225
Skewness			1.000	-0.0259	-0.1398	-0.0623	-0.1059
Kurtosis				1.000	0.0941	-0.0576	-0.0198
SV					1.000	-0.1475	-0.0032
ES(75%)						1.000	0.9726
VaR(75%)							1.000

Table 4.2: Pearson correlation matrix of the risk-premium

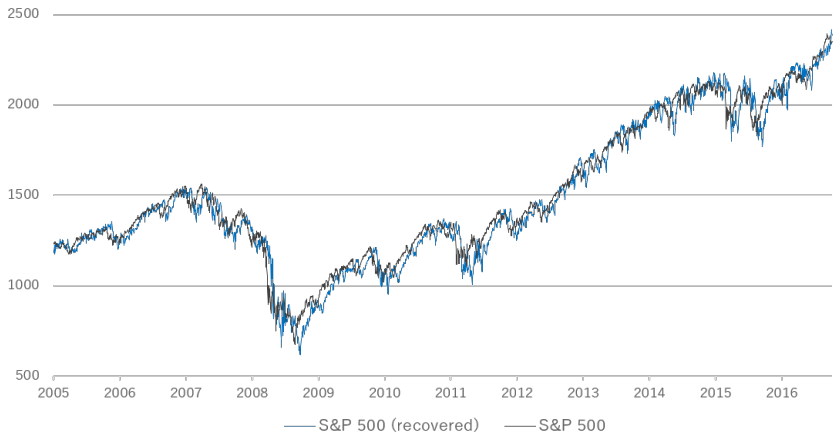


Figure 4.10: Estimated S&P 500 calculated from mean of the recovered physical distribution and actual S&P 500

correlation between the mean and the actual rate of return with lags.

We are predicting the physical distribution 1month(20 business days) later by using the Recovery Theorem. Therefore, we expected the correlation's peak would mark around 20. However, the correlation marks the peak around lag 0 or before 0 in Figure 4.11. It can be interpreted as the physical distribution is influenced mainly by current or past market conditions. Therefore, it means, at least 1st moment, does not have a power to predict the future, and most market participants predict the mean under the physical distribution based on the current or historical movements like VaR.

In the next step, we create EWIs candidates from factors we calculated and investigate their power.

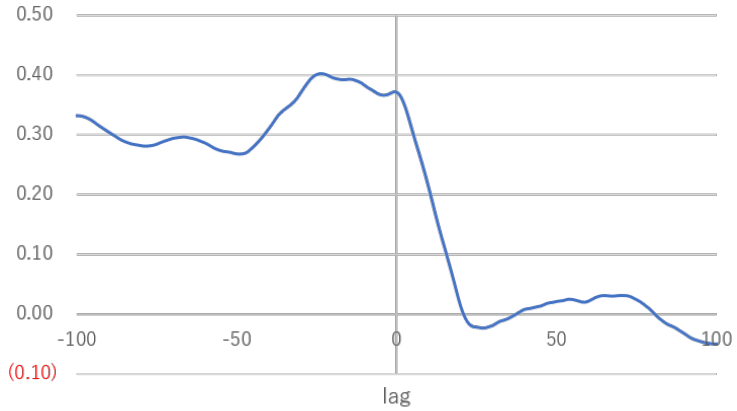


Figure 4.11: Correlation between the recovered physical mean and the actual rate of return with lags

Effectiveness of EWI

First, we calculate the historical average of measure 1 (Table 4.3). Through the backtesting, we use it as the benchmark for checking each EWI power.

Risk occurrence probability	Numerator	Denominator
4.69%	140	2,985

Table 4.3: The risk event occurrence probability (historical average)

We then create 261 EWIs based on the concepts. Table 4.4 shows the top 8 EWIs which mark the highest probabilities in measure 1. $\mathbb{P}$ means physical-measure and $\mathbb{P} - \mathbb{Q}$ means risk-premium, respectively. In addition, Table 4.5 shows the probability of each EWI based on the measure 2.

No	EWI	Measure1	Numerator	Denominator
1	$\mathbb{P}-\mathbb{Q}$-ES(75%)(3days average) exceeds 90%tile	9.8%	14	143
2	$\mathbb{P}$-ES(75%)(3days average) exceeds 90%tile	9.0%	12	133
3	$\mathbb{P}-\mathbb{Q}$-ES(75%)(2days average) exceeds 90%tile	9.0%	17	189
4	$\mathbb{P}$-mean(2days average) exceeds 90%tile	8.8%	20	226
5	$\mathbb{P}-\mathbb{Q}$-ES(90%)(3days average) exceeds 90%tile	8.3%	15	180
6	$\mathbb{P}$-VaR(90%)(2days average)exceeds 90%tile	8.2%	22	269
7	$\mathbb{P}$-ES(75%)(2days average) exceeds 90%tile	8.2%	16	196
8	$\mathbb{P}-\mathbb{Q}$-ES(75%)(1day average) exceeds 90%tile	8.0%	19	238

Table 4.4: Measure 1 of single EWIs chosen by measure 1

No	EWI	Measure2	Numerator	Denominator
1	$\mathbb{P}-\mathbb{Q}$-ES(75%)(3days average) exceeds 90%tile	11.2%	14	125
2	$\mathbb{P}$-ES(75%)(3days average) exceeds 90%tile	9.6%	12	125
3	$\mathbb{P}-\mathbb{Q}$-ES(75%)(2days average) exceeds 90%tile	13.6%	17	125
4	$\mathbb{P}$-mean(2days average) exceeds 90%tile	16.0%	20	125
5	$\mathbb{P}-\mathbb{Q}$-ES(90%)(3days average) exceeds 90%tile	12.0%	15	125
6	$\mathbb{P}$-VaR(90%)(2days average)exceeds 90%tile	17.6%	22	125
7	$\mathbb{P}$-ES(75%)(2days average) exceeds 90%tile	12.8%	16	125
8	$\mathbb{P}-\mathbb{Q}$-ES(75%)(1day average) exceeds 90%tile	15.2%	19	125

Table 4.5: Measure 2 of single EWIs chosen by measure 1

EWIs show, by comparing with the historical average, the risk event occurrence probabilities increase. So it means that recovered distributions may contain some important information for predicting the serious risk events. However, these probabilities are not high enough. Therefore, we investigate the case of 2 EWI combinations.

A slight probability increase is observable in the case of 2 EWI combinations (Table 4.6 and Table 4.7). However, like the result of single EWIs, they still remain low. If we had used other methods such as machine learning, we might have been able to find another EWI that shows the higher probability. Furthermore, we should have removed some pairs which have a high correlation. We treat them as one of the challenges for the future.

EWI	Measure1	Numerator	Denominator
· $\mathbb{P}-\mathbb{Q}$-ES(90%)(2days average) exceeds 90%tile · $\mathbb{P}-\mathbb{Q}$-VaR(90%)(2days average) exceeds 90%tile	11.6%	17	147
· $\mathbb{P}$-VaR(90%)(2days average) exceeds 90%tile · $\mathbb{P}-\mathbb{Q}$-ES(75%)(1day average) exceeds 90%tile	11.1%	15	135
· $\mathbb{P}-\mathbb{Q}$-VaR(75%)(1day average) exceeds 90%tile · $\mathbb{P}-\mathbb{Q}$-ES(75%)(2days average) exceeds 90%tile	10.3%	16	156
· $\mathbb{P}$-ES(75%)(2days average) exceeds 90%tile · $\mathbb{P}-\mathbb{Q}$-variance(2days average) exceeds 90%tile	10.3%	15	146
· $\mathbb{P}-\mathbb{Q}$-ES(75%)(3days average) exceeds 75%tile · $\mathbb{P}$-mean(2days average) exceeds 90%tile	10.2%	16	157
· $\mathbb{P}$-variance(1day average) exceeds 90%tile · $\mathbb{P}$-ES(75%)(2days average) exceeds 90%tile	10.1%	15	149

Table 4.6: Measure 1 of double EWIs chosen by measure 1

EWI	Measure1	Numerator	Denominator
· $\mathbb{P}-\mathbb{Q}$-ES(90%)(2days average) exceeds 90%tile · $\mathbb{P}-\mathbb{Q}$-VaR(90%)(2days average) exceeds 90%tile	13.6%	17	125
· $\mathbb{P}$-VaR(90%)(2days average) exceeds 90%tile · $\mathbb{P}-\mathbb{Q}$-ES(75%)(1day average) exceeds 90%tile	12.0%	15	125
· $\mathbb{P}-\mathbb{Q}$-VaR(75%)(1day average) exceeds 90%tile · $\mathbb{P}-\mathbb{Q}$-ES(75%)(2days average) exceeds 90%tile	12.8%	16	125
· $\mathbb{P}$-ES(75%)(2days average) exceeds 90%tile · $\mathbb{P}-\mathbb{Q}$-variance(2days average) exceeds 90%tile	12.0%	15	125
· $\mathbb{P}-\mathbb{Q}$-ES(75%)(3days average) exceeds 75%tile · $\mathbb{P}$-mean(2days average) exceeds 90%tile	12.8%	16	125
· $\mathbb{P}$-variance(1day average) exceeds 90%tile · $\mathbb{P}$-ES(75%)(2days average) exceeds 90%tile	12.0%	15	125

Table 4.7: Measure 2 of double EWIs chosen by measure 1

4.2 Analysis of USDJPY

In this section, we consider the foreign exchange rate and apply the Recovery Theorem to it. Major currencies which have the higher transaction volume on April 2016 are as below.

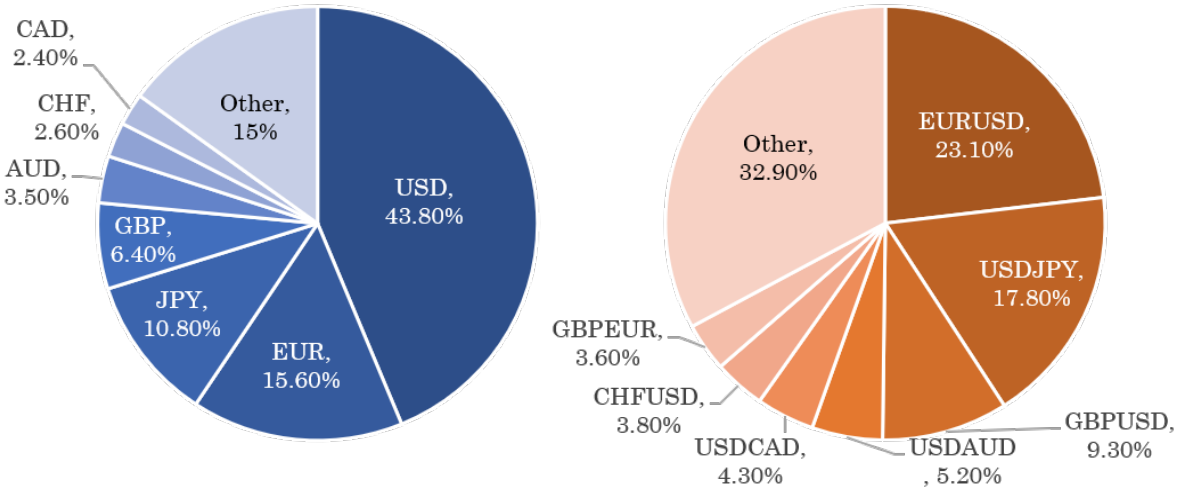


Figure 4.12: Share of trading volume by currencies and currency pairs in April 2016

USD, EUR, JPY and GBP account for over 75% of all trading volume. In the case of the currency pair, USD, as a base currency, is used over 60% of all trading. In this section, we chose USDJPY as the representative of currency exchange rates and apply the same procedure and the same approach (the trinomial tree approach ($k = 4$)) for searching effective EWIs.

4.2.1 Setting

We explain the USDJPY historical data and the risk event used in the analysis. Except them, all methods are the same as the analysis of S&P 500.

Data

Iimplied-volatilities of the foreign currency options are generally stored in Risk-Reversal and Butterfly basis or delta-basis. Figure 4.13 is the implied-volatility screen shot of USDJPY European-option on December 19th, 2017. Actually, they are quoted in Risk-Reversal and Butterfly basis. Therefore, some steps are necessary to get moneyness-basis data.

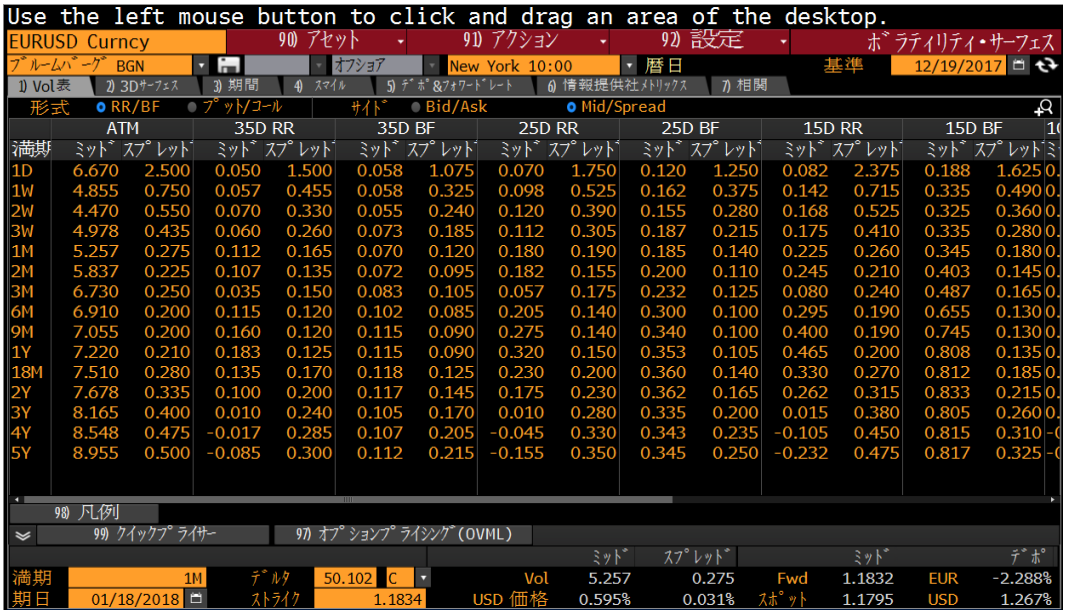


Figure 4.13: Volatility surface for USDJPY on December 19th, 2017 (Data source: Bloomberg L.P.)

The data covers $\{Call_{5\Delta}, Call_{10\Delta}, Call_{15\Delta}, Call_{25\Delta}, Call_{35\Delta}, ATM-delta, Put_{35\Delta}, Put_{25\Delta}, Put_{15\Delta}, Put_{10\Delta}, Put_{5\Delta}\}$ and it is available from June 24th, 2005. After downloading the data from Bloomberg L.P., by solving an optimization problem, we create moneyness-basis implied volatility data[1]. The number of historical data in this analysis is 3,201. Also, 5 option maturities, {30d, 60d, 90d, 180d, 360d}, are used in this analysis.

Risk Event

As Figure 4.2 shows, the risk within 10 business days is designated as the liquidity horizon for USDJPY in BCBS. As we mentioned, it is considered generally BCBS rule is very conservative. In addition, USDJPY marks the second largest amount of volume in the trading. Therefore, we set cumulative 5 business days as the liquidity horizon. Also, for most American companies, increasing USDJPY negatively effects their business. In this case, we assume that we are an American financial institution. So we set over 2% increase in cumulative 5 days within 20 business days" as the risk event.

[1]For more detail about the treatment of foreign exchange rates, see Appendix B

4.2.2 Backtesting

Correlation Between the Actual Return and the Recovered Return

Just like S&P 500, we confirm the correlation between the mean and the actual change ratio of USDJPY with lags (Figure 4.14).

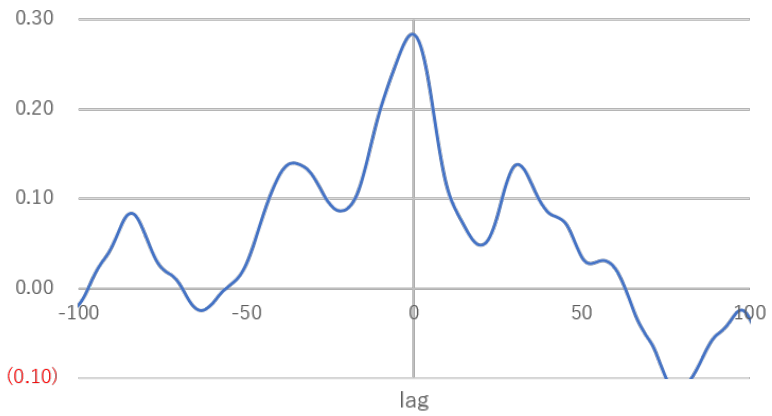


Figure 4.14: Correlation between the recovered physical mean and the actual change ratio of USDJPY with lags

Though the shape of the graph is different from the result of S&P 500, the correlation marks the peak around lag 0 as well. It can be interpreted as the physical distribution is influenced heavily by the current market condition.

Effectiveness of EWI

Based on the 2 concepts described before, we create 261 EWI candidates and investigate their effectiveness by using 2 measures that we already explained.

Risk occurrence probability	Numerator	Denominator
5.5%	176	3,181

Table 4.8: The risk event occurrence probability (historical average)

The result is very similar to S&P 500. Single EWIs show, by comparing with the historical average (Table 4.8), the risk event occurrence probability increases. Also, the fact that the skewness is included all EWIs is an interesting point. So, like S&P 500, it

No	EWI	Measure1	Numerator	Denominator
1	$\mathbb{P}-\mathbb{Q}$-skewness(1days average) exceeds 90%tile	10.4%	22	212
2	$\mathbb{P}$-skewness(1days average)exceeds 90%tile	9.9%	22	222
3	$\mathbb{P}$-skewness(3days average) exceeds 90%tile	9.4%	23	245
4	$\mathbb{P}$-skewness(1days average) exceeds 90%tile	9.0%	38	422
5	$\mathbb{P}$-skewness(90%)(1days average) exceeds 90%tile	9.0%	17	189
6	$\mathbb{P}-\mathbb{Q}$-skewness(2day average) exceeds 90%tile	8.7%	27	310
7	$\mathbb{P}-\mathbb{Q}$-skewness(1days average) exceeds 90%tile	8.5%	36	423
8	$\mathbb{P}-\mathbb{Q}$-skewness(2days average) exceeds 90%tile	8.4%	25	299

Table 4.9: Measure 1 of single EWIs chosen by measure 1

No	EWI	Measure1	Numerator	Denominator
1	$\mathbb{P}-\mathbb{Q}$-skewness(1days average) exceeds 90%tile	12.5%	22	176
2	$\mathbb{P}$-skewness(1days average)exceeds 90%tile	12.5%	22	176
3	$\mathbb{P}$-skewness(3days average) exceeds 90%tile	13.1%	23	176
4	$\mathbb{P}$-skewness(1days average) exceeds 90%tile	21.6%	38	176
5	$\mathbb{P}$-skewness(1days average) exceeds 90%tile	9.7%	17	176
6	$\mathbb{P}-\mathbb{Q}$-skewness(2day average) exceeds 90%tile	15.3%	27	176
7	$\mathbb{P}-\mathbb{Q}$-skewness(1days average) exceeds 90%tile	20.5%	36	176
8	$\mathbb{P}-\mathbb{Q}$-skewness(2days average) exceeds 90%tile	14.2%	25	176

Table 4.10: Measure 2 of single EWIs chosen by measure 1

can be said that recovered distributions may contain some important information for predicting serious risk events. However, these probabilities are not high enough.

The probability increases slightly in the case of 2 EWI combinations. However, probabilities remain very low.

EWI	Measure1	Numerator	Denominator
· $\mathbb{P}$-skewness(1days average) exceeds 90%tile · $\mathbb{P}-\mathbb{Q}$ kurtosis(2days average) exceeds 90%tile	12.4%	19	153
· $\mathbb{P}$-skewness(1days average) exceeds 75%tile · $\mathbb{P}-\mathbb{Q}$ kurtosis(2days average) exceeds 90%tile	12.4%	22	178
· $\mathbb{P}$-skewness(3day average) exceeds 75%tile · $\mathbb{P}-\mathbb{Q}$ skewness(1days average) exceeds 75%tile	11.4%	17	149
· $\mathbb{P}$-skewness(1days average) exceeds 90%tile · $\mathbb{P}-\mathbb{Q}$ skewness(1days average) exceeds 75%tile	11.2%	20	179
· $\mathbb{P}-\mathbb{Q}$ skewness(1days average) exceeds 75%tile · $\mathbb{P}$-skewness(1days average) exceeds 75%tile	10.7%	32	300
· $\mathbb{P}$-skewness(2day average) exceeds 75%tile · $\mathbb{P}-\mathbb{Q}$ skewness(1days average) exceeds 75%tile	10.5%	20	191

Table 4.11: Measure 1 of double EWIs chosen by measure 1

EWI	Measure1	Numerator	Denominator
· $\mathbb{P}$-skewness(1days average) exceeds 90%tile · $\mathbb{P}$-$\mathbb{Q}$ kurtosis(2days average) exceeds 90%tile	10.8%	19	176
· $\mathbb{P}$-skewness(1days average) exceeds 75%tile · $\mathbb{P}$-$\mathbb{Q}$ kurtosis(2days average) exceeds 90%tile	12.5%	22	176
· $\mathbb{P}$-skewness(3day average) exceeds 75%tile · $\mathbb{P}$-$\mathbb{Q}$ skewness(1days average) exceeds 75%tile	9.7%	17	176
· $\mathbb{P}$-skewness(1days average) exceeds 90%tile · $\mathbb{P}$-$\mathbb{Q}$ skewness(1days average) exceeds 75%tile	11.4%	20	176
· $\mathbb{P}$-$\mathbb{Q}$ skewness(1days average) exceeds 75%tile · $\mathbb{P}$-skewness(1days average) exceeds 75%tile	18.2%	32	176
· $\mathbb{P}$-skewness(2day average) exceeds 75%tile · $\mathbb{P}$-$\mathbb{Q}$ skewness(1days average) exceeds 75%tile	11.4%	20	176

Table 4.12: Measure 2 of double EWIs chosen by measure 1

Chapter 5

Concluding Remarks

Ross (2015) has shown that real-world distributions can be derived from risk-neutral densities. However, it is not easy to apply it to real data because it is necessary to solve an ill-posed problem in the process. In this thesis, we propose a new approach, the tree approach, to cope with the problem, and we also apply it to risk management. To the best of our knowledge, this is the only research that applies the Recovery Theorem to risk management.

We would like clarify following 2 points once again. The first is about the new approach. The concept of the tree approach is completely different from other approaches. It has robust theoretical background and concept is easy to understand, just describing transition matrix by using a trinomial tree structure. Also, we show the high accuracy and fast computation time of the recovery under the tree approach. This approach does not need any historical information, so it can be said that it is a better forward-looking approach. In addition, we consider the tree approach with jumps and the non-stationary tree approach. We still can not decide which one is the best approach. However these 2 approaches actually show better results in some specific cases.

The second is the application to risk management. We apply the theorem to real market historical data, S&P 500 and USDJPY. We show moments under the risk-neutral measure are much more stable than the physical distribution measure, and at least 1st moment seems not to have a power for predicting the future physical distribution. We create some EWI candidates and investigate their effectiveness, and actually some of them show better results. Though their power for predicting the future is still low, there is room for improvement.

In this thesis, we create some EWIs and show their effectiveness. However, they do not have power enough to implement into the real business situation, unfortunately.

So, in order to search more effective EWIs, it might be better to use another analytic method. We think deep learning is one of the most useful methods in this case. Although it is difficult to understand the meaning of the factors induced by it, once we develop its program, we can easily find the best factors anytime. Also Overfitting problem happens very often when we use such method. But it might make our analysis much easier to check the most effective indicators before making EWI candidates. Furthermore, deep learning and other machine learning methods are becoming very popular among financial institutions because of the movement of "Fintech". Therefore, from the viewpoint, I think applying deep learning to financial risk management will be very interesting in this topic too.

Second, applying the Recovery Theorem to other financial instruments such as the bond market or other currency pairs is also our remaining task. Martin and Ross (2013) apply the Recovery Theorem to long-term bond market. Hence, we must be able to apply the theorem through the tree approach for the prediction of any interest rates too. Recently, some central banks implemented negative interest policy and it became the most remarkable thing for financial institutions that when the central banks remove the financial policy. So the recovery theorem might be a good predicting tool for searching the time. In the case of foreign exchange rates, Morikawa (2016) applies the theorem to currency pairs for investment strategies. In this thesis, we show only the result of USDJPY. Therefore, another currency pair might have more interesting characteristics.

Also, searching the time of the FX rate regime switching by using the Recovery Theorem might match market practitioners' demands more than simply just predicting the level of the rate. It is generally quite difficult in the FX market to predict the level of the FX rate, because there are so many factors which influence FX rates, and its data includes so much noise. Therefore, predicting the time of the trend might be much easier, and we might be able to get more reliable results.

Appendix A

Breeden–Litzenberger Analysis

In this appendix, we show how to get the state price from the implied volatility data.

Suppose that we have a continuum of prices available for call options with strike K (all with the same time to expiry T). It was originally shown in Breeden and Litzenberger (1978).
We know that a call price is

$$\begin{aligned} C(K,T) &= e^{-rT}\mathbb{E}^{\mathbb{Q}}[(S_T-K)_+], && \text{(A.1)} \\ &= e^{-rT}\int_0^{+\infty}(s-K)_+\pi^*(s)ds, && \text{(A.2)} \\ &= e^{-rT}\int_K^{+\infty}(s-K)\pi^*(s)ds, && \text{(A.3)} \end{aligned}$$

where $\pi^*(s)$ is the probability distribution function for s under the risk-neutral measure.

Taking the first derivative with respect to K follows by differentiating under the integral sign. The result we use is the following. Let $F(x)$ be defined by the following.

$$F(x)=\int_{a(x)}^{b(x)} f(x,s)ds. \tag{A.4}$$

We then have

$$\frac{d}{dx}F(x)=f(x,b(x))b'(x)-f(x,a(x))a'(x)+\int_{a(x)}^{b(x)}\frac{\partial}{\partial x}f(x,s)ds. \tag{A.5}$$

For the Breeden–Litzenberger result, we need to compute (replacing K with x and

forgetting about the discount factor for now)

$$\frac{d}{dx}\int_x^{+\infty}(s-x)\pi^*(s)ds, \tag{A.6}$$

which is equivalent to $(d/dx)F(x)$ as in (A.4) with $a(x)=x, b(x)=+\infty$ and $f(x,s)=(s-x)\pi^*(s)$, noting that $(\partial/\partial x)f(x,s)=-\pi^*(s)$.

From (A.5), we therefore have

$$\begin{aligned}
\frac{d}{dx}F(x) &= -f(x,a(x))a'(x)+\int_{a(x)}^{b(x)}\frac{\partial}{\partial x}f(x,s)ds, &\text{(A.7)}\\
&= -f(x,x)+\int_x^{+\infty}\frac{\partial}{\partial x}f(x,s)ds, &\text{(A.8)}\\
&= \int_x^{+\infty}\frac{\partial}{\partial x}f(x,s)ds, &\text{(A.9)}\\
&= -\int_x^{+\infty}\pi^*(s)ds. &\text{(A.10)}
\end{aligned}$$

A second differentiation under the integral sign is as follows:

$$\begin{aligned}
\frac{d^2}{dx^2}F(x) &= -\frac{d}{dx}\int_x^{+\infty}\pi^*(s)ds, &\text{(A.11)}\\
&= \pi^*(x). &\text{(A.12)}
\end{aligned}$$

From (A.3), we have $F(K)=\int_K^{+\infty}(s-K)\pi^*(s)ds = e^{rT}C(K,T)$ and therefore, by differentiating twice with respect to K and using (A.12), we have the $Breeden-Litzenberger formula$

$$\pi^*(K)=e^{rT}\frac{\partial^2 C(K,T)}{\partial K^2}. \tag{A.13}$$

■

Appendix B

Characteristics of FX Option Data

In this appendix, we first explain the characteristics of the FX option data and Garman–Kohlhagen formula that is usually used to calculate the option implied-volatility. We then show the procedure of making the FX option data on the moneyness-basis.

B.1 Quote Style of the FX Option

The FX option data is available on any market information venders like Bloomberg L.P. Generally, the implied-volatilities are used to show the level of each contract option price.

As Figure 4.13 shows, Risk-Reversal and Butterfly are generally used when the volatilities are quoted on the screen. On the Bloomberg screen, except ATM implied-volatilities, it contains 10 delta and 25 delta Risk-Reversal($RR_{10\Delta}, RR_{25\Delta}$), and 10 delta and 25 delta Butterfly($BF_{10\Delta}, BF_{25\Delta}$). In order to define the finite states for applying the Recovery Theorem, we need to change such data to the data on the moneyness-basis. Risk-Reversal and Butterfly are defined as

$$\begin{aligned} RR &= \sigma_{OTMcall} - \sigma_{OTMput}, && \text{(B.1)} \\ BF &= 0.5Strangle - Straddle, && \text{(B.2)} \\ &= 0.5(\sigma_{OTMcall} + \sigma_{OTMput}) - \sigma_{ATM}. && \text{(B.3)} \end{aligned}$$

Therefore, 4 types of implied-volatilities on the delta-basis are derived from above

relations.

$$\begin{aligned}
\sigma_{call10\Delta} &= \sigma_{ATM} + 0.5RR_{10\Delta} + BF_{10\Delta} && \text{(B.4)}\\
\sigma_{put10\Delta} &= \sigma_{ATM} - 0.5RR_{10\Delta} + BF_{10\Delta} && \text{(B.5)}\\
\sigma_{call25\Delta} &= \sigma_{ATM} + 0.5RR_{25\Delta} + BF_{25\Delta} && \text{(B.6)}\\
\sigma_{put25\Delta} &= \sigma_{ATM} - 0.5RR_{25\Delta} + BF_{25\Delta} && \text{(B.7)}
\end{aligned}$$

B.2 Garman–Kohlhagen Formula

Before the procedure of the implied-volatility data on the moneyness-basis, we explain the Garman–Kohlhagen Formula. It is usually used to calculate option implied-volatilities.

Garman and Kohlhagen (1983) show the European call price formula as bellow under the assumption that FX spot rate $S(t)$ is a random variable with constant volatility σ. Also the risk-free rates of the domestic and foreign currency are defined as r^d and r^f, respectively.

$$\begin{aligned}
\mathbb{E}^{\mathbb{Q}}[e^{-r^dT}(S(T)-K)_+] &= \mathbb{E}^{\mathbb{Q}}\left[e^{-r^dT}\left(S(0)exp\left\{-\sigma\sqrt{T}Y + \left(r - r^f - \frac{1}{2}\sigma^2\right)T\right\} - K\right)_+\right],\\
&= e^{-r^fT}S(0)N(d_1) - e^{-r^dT}KN(d_2), && \text{(B.8)}
\end{aligned}$$

$$d_{1,2} = \frac{1}{\sigma\sqrt{T}}\left[log\frac{S(0)}{K} + \left(r^d - r^f \pm \frac{1}{2}\sigma^2\right)T\right],$$

where Y is a standard normal random variable under risk-neutral measure $\mathbb{Q}$.

Proof

Suppose FX rate $S(t)$, which is the values in the domestic currency to 1 foreign currency, is defined by the following stochastic differential equation under the physical measure. [1]

$$\begin{aligned}
dS(t) &= \gamma(t)S(t)dt + \sigma(t)S(t)\left\{\rho(t)dz_1(t) + \sqrt{1-\rho^2(t)}dz_2(t)\right\}, && \text{(B.9)}\\
&= \gamma(t)S(t)dt + \sigma(t)S(t)dz_3(t), \quad (*Levy\ Theory) && \text{(B.10)}
\end{aligned}$$

where dz_1, dz_2 and dz_3 are Brownian motions, which there is a correlation ρ between

[1] Levy Theory is the theory as below.
Suppose $M(t), t \geq 0$ is a martingale to $\mathcal{F}(t), t \geq 0$ and $M(t)$ satisfies $M(0) = 0$. Also suppose $M(t)$ has continuous path and $[M, M](t) = t$ for all $t \geq 0$. Under such situation, $M(t)$ is a Brownian motion.

dz_1 and dz_2. Think about investing in the foreign money-market-account and translating the money into the domestic currency. Foreign money-market-account value on the domestic currency basis is defined as $M^f(t)S(t)$, where $M^f(t)$ is a foreign money-market-account, and we define the discount factor as $D(t)$. Also, we define the foreign risk-free-rate as $r^f(t)$ and similarly domestic risk-free-rate as $r^d(t)$. Hence, the stochastic differential equation of $D(t)M^f(t)S(t)$ can be described as bellow:

$$d(D(t)M^f(t)S(t)) = D(t)M^f(t)S(t)[(r^f(t) - r^d(t) + \gamma(t))dt + \sigma(t)dz_3]. \tag{B.11}$$

In this case, since we can calculate an unique market price of risk, we can chose a unique risk-neutral distribution $\mathbb{Q}$.

$$d(D(t)M^f(t)S(t)) = D(t)M^f(t)S(t)[\sigma(t)dz_3^{\mathbb{Q}}(t)]. \tag{B.12}$$

Multiplying $M(t) = \frac{1}{D(t)}$, $D^f(t)$ and (B.12), then we get

$$dS(t) = S(t)[(r^d(t) - r^f(t))dt + \sigma(t)dz_3^{\mathbb{Q}}(t)]. \tag{B.13}$$

If r^d,r^f and σ are assumed as fixed numbers, we get

$$S(T) = S(t)exp\left\{\left(r^d - r^f - \frac{1}{2}\sigma^2\right)T + \sigma dz_3^{\mathbb{Q}}\right\}. \tag{B.14}$$

Since the payoff of the European call option on the domestic-basis is $(S(T) - K)_+$, we therefore can describe the time-t value $V_{d/f}$ as

$$V_{d/f} \equiv \mathbb{E}^{\mathbb{Q}}[e^{-r^dT}(S(T) - K)_+]. \tag{B.15}$$

Putting (B.14) into (B.15), we have

$$V_{d/f} = \mathbb{E}^{\mathbb{Q}}\left[e^{-r^dT}\left(S(0)exp\left\{-\sigma\sqrt{T}Y + \left(r^d - r^f - \frac{1}{2}\sigma^2\right)T\right\} - K\right)_+\right]. \tag{B.16}$$

Therefore, like Black-Scholes formula, we can calculate the call option present value as follows:

$$\begin{aligned} V_{d/f} &= e^{-r^fT}S(0)N(d_1) - e^{-r^dT}KN(d_2), \\ d_{1,2} &= \frac{1}{\sigma\sqrt{T}}\left[log\frac{S(0)}{K} + \left(r^d - r^f \pm \frac{1}{2}\sigma^2\right)T\right]. \end{aligned} \tag{B.17}$$

■

B.3 FX Delta

In this section, we explain the quote style convention. In currency markets, as opposed to equity markets, options can be quoted in one of four relative quote styles, domestic per foreign(d/f)(or $d;pips$), percentage foreign($\%f$), percentage domestic(%d) and foreign per domestic(f/d)(or $f;pips$). This is because, unlike equities, investors can have two numeraires. A risk-neutral investor in the domestic currency can therefore obtain a domestic per domestic price or a domestic per foreign price. Similarly, a risk-neutral investor in the foreign currency can obtain a foreign per domestic price or a foreign per foreign price. European option prices based on the quote style are summarized as below:

$$V_{d;pips} \equiv V_{d/f} = \omega S(0) e^{-r^f T} N(\omega d_1) - \omega K e^{-r^d T} N(\omega d_2), \tag{B.18}$$

$$V_{f\%} = \frac{V_{d/f}}{S_0}, \tag{B.19}$$

$$V_{d\%} = \frac{V_{d/f}}{K}, \tag{B.20}$$

$$V_{f;pips} \equiv V_{f/d} = \frac{V_{d/f}}{S_0 K}, \tag{B.21}$$

where ω is 1 in the case of the call option, and in the case of put option, ω is -1.

Quote styles of EURUSD, USDJPY and GBPUSD are as follows.

Currency pair	Base currency	Quote currency	Premium	Quote style
EURUSD	EUR	USD	USD	f;pips
USDJPY	USD	JPY	USD	f %
GBPUSD	GBP	USD	USD	f;pips

Table B.1: Quote style

"Base currency"and "Quote currency" are currencies which satisfies $\left(FX\ rate = \frac{Quote\ currency}{Base\ currency}\right)$. Also, "premium" is the option premium currency. Therefore, when the FX option price is quoted on the screen, the price on the premium-currency-basis is used.

In USDJPY, the quote style of this pair is defined as $V_{f,\%}$. We then explain how to calculate the delta from European option prices.

European call % delta $\Delta_{C;\%}$ is calculated as

$$\begin{aligned}
\Delta_{C;\%} &= \lim_{\Delta S \to 0} \frac{\Delta V_{f,\%}}{\Delta S/S}, && \text{(B.22)}\\
&= \frac{\partial (V_{d,pips}/S)}{\Delta S/S}, && \text{(B.23)}\\
&= S\frac{\partial \left(V_{d,pips}/S\right)}{\partial S}, && \text{(B.24)}\\
&= S\frac{(\partial V_{d,pips}/\partial S)S - V_{d,pips}}{S^2}, && \text{(B.25)}\\
&= \frac{\partial V_{d,pips}}{\partial S} - \frac{V_{d,pips}}{S}, && \text{(B.26)}\\
&= \frac{\partial (V_{d,pips})}{\partial S} - \frac{e^{-r^f T}SN(d_1) - Ke^{-r^d T}N(d_2)}{S}. && \text{(B.27)}
\end{aligned}$$

Then, we use

$$n(d_1) = n(d_2)exp\{(r^f - r^d)T\}(K/S), \tag{B.28}$$

$$\begin{aligned}
\frac{\partial V_{d,pips}}{\partial S} &= \frac{\partial \{e^{-r^f T}SN(d_1) - Ke^{-r^d T}N(d_2)\}}{\partial S}, && \text{(B.29)}\\
&= e^{-r^f T}N(d_1) + Se^{-r^f T}\frac{\partial N(d_1)}{\partial S} - Ke^{-r^d T}\frac{\partial N(d_2)}{\partial S}, && \text{(B.30)}\\
&= e^{-r^f T}N(d_1) + \frac{1}{\sigma S\sqrt{T}}\left(Se^{-r^f T}n(d_1) - Ke^{-r^d T}n(d_2)\right), && \text{(B.31)}\\
&= e^{-r^f T}N(d_1). && \text{(B.32)}
\end{aligned}$$

Therefore, $\Delta_{C;\%}$ is expressed as below:

$$\begin{aligned}
\Delta_{C;\%} &= e^{-r_f T}N(d_1) - e^{-r_f T}N(d_1) + \frac{Ke^{-r_d T}N(d_2)}{S}, && \text{(B.33)}\\
&= e^{-r_d T}\frac{K}{S}N(d_2). && \text{(B.34)}
\end{aligned}$$

Also, apply same approach to calculate European put % delta $\Delta_{P;\%}$. According to Garman and Kohlhagen (1983), the European put option price is described as

$$\begin{aligned}
\mathbb{E}^{\mathbb{Q}}[e^{-r^d T}(K - S(T))_+] &= -e^{-r^f T}S(0)N(-d_+) + e^{-r^d T}KN(-d_-), && \text{(B.35)}\\
& d_\pm = \frac{1}{\sigma\sqrt{T}}\left[log\frac{S(0)}{K} + \left(r - r^f \pm \frac{1}{2}\sigma^2\right)T\right].
\end{aligned}$$

So we simply calculate $\Delta_{P;\%}$ like $\Delta_{C;\%}$.

$$\begin{aligned}
\Delta_{P;\%} &= \lim_{\Delta S \to 0} \frac{\Delta V_{f,\%}}{\Delta S/S}, && \text{(B.36)}\\
&= \frac{\partial V_{d,pips}}{\partial S} - \frac{V_{d,pips}}{S}, && \text{(B.37)}\\
&= \frac{\partial (V_{d,pips})}{\partial S} - \frac{-e^{-r^f T} S N(-d_1) + K e^{-r^d T} N(-d_2)}{S}. && \text{(B.38)}
\end{aligned}$$

The first term in (B.38) can be expressed as

$$\begin{aligned}
\frac{\partial V_{d,pips}}{\partial S} &= \frac{\partial \{-e^{-r^f T} S N(-d_1) + K e^{-r^d T} N(-d_2)\}}{\partial S}, && \text{(B.39)}\\
&= -e^{-r^f T} N(-d_1) - S e^{-r^f T} \frac{\partial N(-d_1)}{\partial S} + K e^{-r^d T} \frac{\partial N(-d_2)}{\partial S}, && \text{(B.40)}\\
&= -e^{-r^f T} N(-d_1) + \frac{1}{\sigma S \sqrt{T}} \left(S e^{-r^f T} n(-d_1) - K e^{-r^d T} n(-d_2) \right), && \text{(B.41)}\\
&= -e^{-r^f T} N(-d_1). && \text{(B.42)}
\end{aligned}$$

Therefore, European put delta $\Delta_{P;\%}$ is described as below:

$$\begin{aligned}
\Delta_{P;\%} &= \frac{\partial (V_{d,pips})}{\partial S} - \frac{-e^{-r^f T} S N(-d_1) + K e^{-r^d T} N(-d_2)}{S}, && \text{(B.43)}\\
&= -e^{-r^d T} \frac{K}{S} N(-d_2). && \text{(B.44)}
\end{aligned}$$

Similarly, we can calculate $\Delta_{C;pips}$ and $\Delta_{P;pips}$ as follows:

$$\begin{aligned}
\Delta_{C;pips} &= e^{-r^f T} N(d_1), && \text{(B.45)}\\
\Delta_{P;pips} &= -e^{-r^f T} N(-d_1). && \text{(B.46)}
\end{aligned}$$

B.4 Procedure of Making the Implied-Volatility on the Moneyness-Basis

Finally, we explain about the procedure of making the implied-volatility data on the moneyness-basis. The process is composed by 4 steps.

Procedure

1. Calculate the option price by using Garman–Kohlhagen formula with respect to an arbitrary implied-volatility IV and strike K.

2. Calculate the delta with respect to each option price.

3. Compare the delta in step 2 with the real-market implied-volatility data IV_{smile}.

4. Chose the optimal K which minimizes the difference between IV and IV_{smile}.

ATM Strike

At last, we explain about the ATM strike data. There are basically two possibilities that are used in practice. First one is straightforward, "at-the-money-strike". For a particular maturity T at time 0, the strike K_{ATM} is set to the froward value $F_{0,T}$.

$$K_{ATM} \equiv F_{0,T} \tag{B.47}$$

However, this style is only used for currency pairs including a Latin American emerging market currency.

A more natural way to define the at-the-money strike is the strike K_{ATM} for which it is possible to buy a straddle that corresponds to a pure long vega position with no net delta. This is known as the "delta-neutral-straddle", and it has the advantage that such a trade can be effected without any spot or forward trade being needed. If we define $\Delta(\{Call, Put\}, K, T, \sigma)$ as the function calculating the delta, the delta-neutral-straddle strike K_{DNS} satisfies the following formula.

$$\Delta(Call, K_{DNS}, T, \sigma_{ATM}) + \Delta(Put, K_{DNS}, T, \sigma_{ATM}) = 0. \tag{B.48}$$

Same as the option premium, there are 2 quote styles, $K_{DNS;\%}$ and $K_{DNS;pips}$.

DNS Strike of % delta ($K_{DNS;\%}$)

We get the following relation by putting (B.34),(B.44) into (B.48)

$$N\left(\frac{lnS_0 - lnK_{DNS;\%} + r^d T - r^f T - \frac{1}{2}\sigma^2 T}{\sigma\sqrt{T}}\right) = N\left(\frac{lnK_{DNS;\%} - lnS_0 - r^d T + r^f T + \frac{1}{2}\sigma^2 T}{\sigma\sqrt{T}}\right).$$

Therefore, we get $K_{DNS;\%}$ as below:

$$\begin{aligned} lnK_{DNS;\%} &= lnS_0 - \frac{1}{2}\sigma^2 T + r^d T - r^f T, & \text{(B.49)} \\ \therefore K_{DNS;\%} &= S_0 exp\left((r^d - r^f)T\right) exp\left(-\frac{1}{2}\sigma^2 T\right). & \text{(B.50)} \end{aligned}$$

DNS Strike of pips delta ($K_{DNS;pips}$)

We similarly get following relation.

$$N\left(\frac{lnS_0 - lnK_{DNS;pips} + r^d T - r^f T + \frac{1}{2}\sigma^2 T}{\sigma\sqrt{T}}\right) = N\left(\frac{lnK_{DNS;pips} - lnS_0 - r^d T + r^f T - \frac{1}{2}\sigma^2 T}{\sigma\sqrt{T}}\right).$$

So, we can describe $K_{DNS;pips}$ as below.

$$lnK_{DNS;pips} = lnS_0 + \frac{1}{2}\sigma^2 T + r^d T - r^f T, \tag{B.51}$$

$$\therefore K_{DNS;pips} = S_0 exp\left((r^d - r^f)T\right) exp\left(\frac{1}{2}\sigma^2 T\right). \tag{B.52}$$

Hence, we easily decide the strike price by using (B.50) and (B.52).

References

[1] Alexander, J.M., R. Frey and P. Embrechts (2008). *Quantitative Risk Management.* Princeton University Press, Princeton.

[2] Audrino, F., R. Huitema and M. Ludwig (2015). "An empirical analysis of the Ross recovery theorem." Working Paper, University of St. Gallen.

[3] Bank for international settlements. (2015). "Guidance on credit risk and accounting for expected credit losses."

[4] Bank for international settlements (2016). "Minimum capital requirements for market risk."

[5] Bliss, R.R. and N. Panigirtzoglou (2002). "Testing the stability of implied probability density functions." *Journal of Banking Finance.* Vol 26. 381–422.

[6] Breeden, D.T. and R.H. Litzenberger (1978). "Prices of state-contingent claims implicit in option prices." *The Journal of Business* Vol 51. 621–651.

[7] Carr, P. and J. Yu (2012). "Risk, return and Ross recovery." *The Journal of Derivatives.* Vol 20. 38–59.

[8] Clark, I.J. (2011). *Foreign exchange option pricing, A practitioner's Guide.* The Wiley Finance Series.

[9] Derman, E. and I. Kani (1994). "The volatility smile and its implied tree." Working Paper, Goldman Sachs & Co.

[10] Derman, E., I. Kani and N. Chriss (1996). "Implied trinomial trees of the volatility smile." Working Paper, Goldman Sachs & Co.

[11] Dubkynskiy, S. and R.S. Goldstein (2013). "Recovering drift and preference parameters from financial derivatives." Working Paper, University of Minnesota.

[12] European banking authority. (2017). "2018 EU-wide stress test - methodological note."

[13] Fabio, M., W. Julian and Z. Yang (2016). "Empirical recovery : Hansen–Scheinkman factorization and Ross recovery from high-frequency option prices." *Journal of Operations Research.*

[14] Federal reserve board. (2017). "Comprehensive capital analysis and review 2017: assessment framework and results."

[15] Bank of England. (2017). "Stress testing the UK banking system: 2017 guidance for participating banks and building societies."

[16] Figlewski, S. (2008). "Estimating the implied risk neutral density for the U.S. market portfolio." Oxford University Press, Oxford.

[17] Garman, M.B. and S.W. Kohlhagen (1983). "Foreign currency option values." *Journal of International Money and Finance.* Vol.2. 231–237.

[18] Jensen, C.S., D. Lando and L.H. Pedersen (2016). "Generalized recovery." Working Paper, Copenhagen Business School.

[19] Kiriu, T. and N. Hibiki (2015). "Estimating forward-looking distribution with the Ross recovery theorem." Working Paper, Keio University.

[20] Kou, S.G. (2002). "Jump diffusion model for option pricing." *Management Science.* Vol 48. 1086–1101.

[21] Ludwig, M. (2015). "Robust estimation of shape-constrained state price density surfaces." *The Journal of Derivatives.* Vol 22. 56–72.

[22] Martin, I. and S.A. Ross (2013). "The log bond." Working Paper, Stanford University.

[23] Merton, R. (1976). "Option pricing when underlying stock return are discontinuous." *Journal of Financial Economics.* Vol 3. 125–144.

[24] Mizuho financial group, Inc. (2017). "Integrated report 2017 (annual review))."

[25] Morikawa, R. (2016). "Ross Recovery on FX market." Master's thesis for Hitotsubashi University. (in Japanese)

[26] Park, H. (2014). "Ross recovery with recurrent and transient processes." *Quantitative Finance.* Vol 16. 667–676.

[27] Ross, S. (2015). "The recovery theorem." *The Journal of Finance.* Vol 70. 615–648.

[28] Shreve S.E. (2003). *Stochastic Calculus for Finance 2: Continuous Time Models.* Springer Finance.

[29] Walden, J. (2014). "Recovery with unbounded diffusion processes." *Review of Finance.* Vol 21. 1403–1444.

[30] William, R.M. and C.P. Thomas (1997). "Recovering an asset's implied pdf from option prices: an application to crude oil during the gulf crisis." *Journal of Financial and Quantitative Analysis.* Vol 32. 91–115.